GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 18, 2004, 12:08:12 ; Search time 48 Seconds
(without alignments)
2104.352 Million cell updates/sec

Title: US-10-049-569-2
Perfect score: 1858
Sequence: 1 MGPGEALLAGLLVMVLAVAL.....ASTHNGSVDTENDSCLQQTH 363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1145568 seqs, 278261457 residues

Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1858	100.0	363	9	US-09-736-131-2	Sequence 2, Appli
2	1858	100.0	363	10	US-09-782-974C-186	Sequence 186, App
3	1858	100.0	363	12	US-10-206-915-590	Sequence 590, App
4	1858	100.0	363	12	US-10-199-670-590	Sequence 590, App
5	1858	100.0	363	12	US-10-201-858-590	Sequence 590, App
6	1858	100.0	363	12	US-10-205-890-590	Sequence 590, App
7	1858	100.0	363	12	US-10-208-024-590	Sequence 590, App
8	1858	100.0	363	12	US-10-201-853-590	Sequence 590, App
9	1858	100.0	363	12	US-10-174-581-590	Sequence 590, App
10	1858	100.0	363	12	US-10-176-483-590	Sequence 590, App
11	1858	100.0	363	12	US-10-176-749-590	Sequence 590, App
12	1858	100.0	363	12	US-10-176-914-590	Sequence 590, App
13	1858	100.0	363	12	US-10-176-915-590	Sequence 590, App
14	1858	100.0	363	12	US-10-176-484-590	Sequence 590, App
15	1858	100.0	363	12	US-10-180-550-590	Sequence 590, App
16	1858	100.0	363	12	US-10-183-014-590	Sequence 590, App
17	1858	100.0	363	12	US-10-187-738-590	Sequence 590, App
18	1858	100.0	363	12	US-10-187-740-590	Sequence 590, App
19	1858	100.0	363	12	US-10-187-883-590	Sequence 590, App
20	1858	100.0	363	12	US-10-194-363-590	Sequence 590, App
21	1858	100.0	363	12	US-10-194-460-590	Sequence 590, App
22	1858	100.0	363	12	US-10-194-463-590	Sequence 590, App
23	1858	100.0	363	12	US-10-194-484-590	Sequence 590, App
24	1858	100.0	363	12	US-10-195-884-590	Sequence 590, App
25	1858	100.0	363	12	US-10-195-896-590	Sequence 590, App
26	1858	100.0	363	12	US-10-196-744-590	Sequence 590, App
27	1858	100.0	363	12	US-10-196-755-590	Sequence 590, App
28	1858	100.0	363	12	US-10-196-757-590	Sequence 590, App
29	1858	100.0	363	12	US-10-197-704-590	Sequence 590, App
30	1858	100.0	363	12	US-10-197-710-590	Sequence 590, App
31	1858	100.0	363	12	US-10-198-758-590	Sequence 590, App
32	1858	100.0	363	12	US-10-198-766-590	Sequence 590, App
33	1858	100.0	363	12	US-10-199-304-590	Sequence 590, App
34	1858	100.0	363	12	US-10-199-309-590	Sequence 590, App
35	1858	100.0	363	12	US-10-199-313-590	Sequence 590, App
36	1858	100.0	363	12	US-10-199-456-590	Sequence 590, App
37	1858	100.0	363	12	US-10-201-329-590	Sequence 590, App
38	1858	100.0	363	12	US-10-202-412-590	Sequence 590, App
39	1858	100.0	363	12	US-10-206-919-590	Sequence 590, App
40	1858	100.0	363	12	US-10-206-922-590	Sequence 590, App
41	1858	100.0	363	12	US-10-206-924-590	Sequence 590, App
42	1858	100.0	363	12	US-10-206-928-590	Sequence 590, App
43	1858	100.0	363	12	US-10-207-914-590	Sequence 590, App
44	1858	100.0	363	12	US-10-207-921-590	Sequence 590, App
45	1858	100.0	363	12	US-10-207-922-590	Sequence 590, App

ALIGNMENTS

RESULT 1
US-09-736-131-2
; Sequence 2, Application US/09736131
; Patent No. US20020019347A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER et al.
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000738
; CURRENT APPLICATION NUMBER: US/09/736,131
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: 60/219,449
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Human
US-09-736-131-2

Query Match 100.0%; Score 1858; DB 9; Length 363;
Best Local Similarity 100.0%; Pred. No. 2.4e-169;
Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
```

```
Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
              |||
Db        361 QTH 363
```

RESULT 2
US-09-782-974C-186
; Sequence 186, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; APPLICANT: Lind, Peter
; APPLICANT: Wood, Linda S.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor
; FILE REFERENCE: 411USPHRM311
; CURRENT APPLICATION NUMBER: US/09/782,974C
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/165,838
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 09/714,449
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 60/198,568
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,071
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,678
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 60/173,396
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/184,129
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/185,421
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,554
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,530
; PRIOR FILING DATE: 2000-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 186
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-782-974C-186

Query Match 100.0%; Score 1858; DB 10; Length 363;
Best Local Similarity 100.0%; Pred. No. 2.4e-169;
Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
              |||
Db        361 QTH 363
```

RESULT 3
US-10-206-915-590
; Sequence 590, Application US/10206915
; Publication No. US20040029221A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C513
; CURRENT APPLICATION NUMBER: US/10/206,915
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 590
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-206-915-590

Query Match 100.0%; Score 1858; DB 12; Length 363;
Best Local Similarity 100.0%; Pred. No. 2.4e-169;
Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
                 |||
Db        361 QTH 363
```

RESULT 4
US-10-199-670-590
; Sequence 590, Application US/10199670
; Publication No. US20040033560A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C401
; CURRENT APPLICATION NUMBER: US/10/199,670
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 590
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-199-670-590

Query Match 100.0%; Score 1858; DB 12; Length 363;
Best Local Similarity 100.0%; Pred. No. 2.4e-169;
Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
                 |||
Db        361 QTH 363
```

RESULT 5
US-10-201-858-590
; Sequence 590, Application US/10201858
; Publication No. US20040038337A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C464
; CURRENT APPLICATION NUMBER: US/10/201,858
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540

; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 590
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-201-858-590

Query Match 100.0%; Score 1858; DB 12; Length 363;
Best Local Similarity 100.0%; Pred. No. 2.4e-169;
Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
              |||
Db        361 QTH 363
```

RESULT 6
US-10-205-890-590
; Sequence 590, Application US/10205890
; Publication No. US20040048334A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C519
; CURRENT APPLICATION NUMBER: US/10/205,890
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 590
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-205-890-590

Query Match 100.0%; Score 1858; DB 12; Length 363;
Best Local Similarity 100.0%; Pred. No. 2.4e-169;
Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
              |||
Db        361 QTH 363
```

RESULT 7
US-10-208-024-590
; Sequence 590, Application US/10208024
; Publication No. US20040048335A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME

```
;  FILE REFERENCE: P3430R1C538
;  CURRENT APPLICATION NUMBER: US/10/208,024
;  CURRENT FILING DATE:  2002-07-29
;  PRIOR APPLICATION NUMBER: 10/052586
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059266
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/063120
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063121
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063486
;  PRIOR FILING DATE: 1997-10-21
;  PRIOR APPLICATION NUMBER: 60/063540
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063541
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063544
;  PRIOR FILING DATE: 1997-10-28
;  Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 590
;   LENGTH: 363
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-208-024-590

  Query Match             100.0%;  Score 1858;  DB 12;  Length 363;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-169;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
              |||
Db        361 QTH 363


RESULT 8
US-10-201-853-590
; Sequence 590, Application US/10201853
; Publication No. US20040053358A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C465
;  CURRENT APPLICATION NUMBER: US/10/201,853
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: 10/052586
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059266
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/063120
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063121
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063486
;  PRIOR FILING DATE: 1997-10-21
;  PRIOR APPLICATION NUMBER: 60/063540
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063541
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063544
;  PRIOR FILING DATE: 1997-10-28
;  Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 590
;   LENGTH: 363
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-201-853-590

  Query Match             100.0%;  Score 1858;  DB 12;  Length 363;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-169;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
              |||
Db        361 QTH 363
```

RESULT 9
US-10-174-581-590
; Sequence 590, Application US/10174581
; Publication No. US20030017540A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C41
; CURRENT APPLICATION NUMBER: US/10/174,581
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086023
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086486
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087098
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087208
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759

```
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/087827
;  PRIOR FILING DATE: 1998-06-03
;  PRIOR APPLICATION NUMBER: 60/088025
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/088028
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/088029
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/088033
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/088167
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/088202
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/088212
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/088217
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/088326
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/088655
;  PRIOR FILING DATE: 1998-06-09
;  PRIOR APPLICATION NUMBER: 60/088722
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088738
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088740
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088811
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088824
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088825
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088826
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088861
;  PRIOR FILING DATE: 1998-06-11
;  PRIOR APPLICATION NUMBER: 60/088863
;  PRIOR FILING DATE: 1998-06-11
;  PRIOR APPLICATION NUMBER: 60/088876
;  PRIOR FILING DATE: 1998-06-11
;  PRIOR APPLICATION NUMBER: 60/089090
;  PRIOR FILING DATE: 1998-06-12
;  PRIOR APPLICATION NUMBER: 60/089105
;  PRIOR FILING DATE: 1998-06-12
;  PRIOR APPLICATION NUMBER: 60/089512
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: 60/089514
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: 60/089538
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089598
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089653

  Query Match            100.0%;  Score 1858;  DB 12;  Length 363;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-169;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
              |||
Db        361 QTH 363


RESULT 10
US-10-176-483-590
; Sequence 590, Application US/10176483
; Publication No. US20030017541A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C68
;  CURRENT APPLICATION NUMBER: US/10/176,483
;  CURRENT FILING DATE:  2002-06-20
;  Prior application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 590
;   LENGTH: 363
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-176-483-590

  Query Match            100.0%;  Score 1858;  DB 12;  Length 363;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-169;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
```

```
Qy        361 QTH 363
              |||
Db        361 QTH 363


RESULT 11
US-10-176-749-590
; Sequence 590, Application US/10176749
; Publication No. US20030017542A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C76
;  CURRENT APPLICATION NUMBER: US/10/176,749
;  CURRENT FILING DATE:  2002-06-20
;  Prior application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 590
;   LENGTH: 363
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-176-749-590

  Query Match             100.0%;  Score 1858;  DB 12;  Length 363;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-169;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
              |||
Db        361 QTH 363


RESULT 12
US-10-176-914-590
; Sequence 590, Application US/10176914
; Publication No. US20030017543A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C83
;  CURRENT APPLICATION NUMBER: US/10/176,914
;  CURRENT FILING DATE:  2002-06-20
;     Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 590
;   LENGTH: 363
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-176-914-590

  Query Match             100.0%;  Score 1858;  DB 12;  Length 363;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-169;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
              |||
Db        361 QTH 363


RESULT 13
US-10-176-915-590
; Sequence 590, Application US/10176915
; Publication No. US20030017544A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
```

```
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C110
;  CURRENT APPLICATION NUMBER: US/10/176,915
;  CURRENT FILING DATE:  2002-06-21
;     Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 590
;    LENGTH: 363
;    TYPE: PRT
;    ORGANISM: Homo Sapien
US-10-176-915-590

  Query Match             100.0%;  Score 1858;  DB 12;  Length 363;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-169;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
              |||
Db        361 QTH 363


RESULT 14
US-10-176-484-590
; Sequence 590, Application US/10176484
; Publication No. US20030059876A9
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C64
;  CURRENT APPLICATION NUMBER: US/10/176,484
;  CURRENT FILING DATE:  2002-06-20
;  Prior application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 590
;    LENGTH: 363
;    TYPE: PRT
;    ORGANISM: Homo Sapien
US-10-176-484-590

  Query Match             100.0%;  Score 1858;  DB 12;  Length 363;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-169;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
              |||
Db        361 QTH 363


RESULT 15
US-10-180-550-590
; Sequence 590, Application US/10180550
; Publication No. US20030064440A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C149
;  CURRENT APPLICATION NUMBER: US/10/180,550
;  CURRENT FILING DATE:  2002-06-25
;  Prior application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 590
;    LENGTH: 363
;    TYPE: PRT
;    ORGANISM: Homo Sapien
US-10-180-550-590

  Query Match             100.0%;  Score 1858;  DB 12;  Length 363;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-169;
  Matches  363;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy         121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy         181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240

Qy         241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy         301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy         361 QTH 363
                |||
Db         361 QTH 363
```

Search completed: May 18, 2004, 12:14:05
Job time : 50 secs

OM protein - protein search, using sw model

Run on: May 18, 2004, 12:04:02 ; Search time 20 Seconds
(without alignments)
1745.876 Million cell updates/sec

Title: US-10-049-569-2
Perfect score: 1858
Sequence: 1 MGPGEALLAGLLVMVLAVAL.....ASTHNGSVDTENDSCLQQTH 363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	248	13.3	437	2	I57942	5-hydroxytryptamin
2	244.5	13.2	440	2	JC5520	serotonin receptor
3	243.5	13.1	388	2	JN0605	somatostatin recep
4	239	12.9	436	2	JN0591	serotonin receptor
5	232.5	12.5	359	2	JC4120	histamine H2 recep
6	231.5	12.5	364	2	JN0763	somatostatin recep
7	231	12.4	363	2	I57955	somatostatin recep
8	230.5	12.4	408	1	QRHUBE	beta-3-adrenergic
9	230.5	12.4	414	1	QRHUB3	beta-3-adrenergic
10	229.5	12.4	418	2	G02953	beta-3-adrenergic
11	228.5	12.3	363	2	I57940	somatostatin recep
12	228.5	12.3	391	2	A41795	somatostatin recep
13	228.5	12.3	391	2	C41795	somatostatin recep
14	228.5	12.3	391	2	A39297	somatostatin recep
15	228.5	12.3	477	1	DYHUD5	dopamine receptor
16	228	12.3	387	2	JC5949	galanin receptor 2
17	228	12.3	475	2	A41271	dopamine receptor
18	226.5	12.2	420	2	T25689	hypothetical prote
19	224	12.1	457	2	I51660	dopamine D1B recep
20	223.5	12.0	359	2	A39008	histamine H2 recep
21	223	12.0	369	2	B41795	somatostatin recep
22	223	12.0	445	2	A48881	serotonin receptor
23	222.5	12.0	384	2	A47249	brain-specific som
24	222	11.9	369	2	A45291	somatostatin recep
25	222	11.9	448	2	A47519	serotonin receptor
26	222	11.9	459	2	A56849	dopamine receptor-
27	221	11.9	369	2	JC2083	somatostatin recep
28	220	11.8	359	2	JH0449	histamine H2 recep
29	220	11.8	370	2	I48231	serotonin receptor
30	220	11.8	477	2	S71323	alpha-1A adrenergi
31	219	11.8	564	2	A38271	serotonin receptor
32	216	11.6	387	1	DYHUD4	dopamine receptor
33	216	11.6	448	2	S36402	serotonin receptor
34	216	11.6	560	2	A38731	alpha-1A adrenergi
35	215.5	11.6	358	2	JQ1278	histamine H2 recep
36	215	11.6	387	2	S55550	5-HT4S receptor -
37	215	11.6	451	2	I51659	dopamine D1A recep
38	214.5	11.5	444	2	C55886	dopamine receptor
39	214	11.5	418	2	A46226	somatostatin recep
40	213.5	11.5	400	2	S32804	beta-3-adrenergic
41	213	11.5	369	2	D41795	somatostatin recep
42	212.5	11.4	363	2	I50475	dopamine D1 recept
43	212	11.4	406	2	S55549	serotonin 4 recept
44	212	11.4	428	2	S30508	probable G protein
45	212	11.4	487	1	DYRTD1	dopamine receptor

ALIGNMENTS

RESULT 1
I57942
5-hydroxytryptamine receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C;Accession: I57942
R;Monsma, F.J.
Mol. Pharmacol. 43, 320-327, 1993
A;Title: Cloning and expression of a novel serotonin receptor with high affinity for tric
A;Reference number: I57942; MUID:93196608; PMID:7680751
A;Accession: I57942
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-437 <RES>
A;Cross-references: GB:L03202; NID:g294507; PIDN:AAA40618.1; PID:g294508
C;Superfamily: vertebrate rhodopsin

Query Match 13.3%; Score 248; DB 2; Length 437;
Best Local Similarity 25.3%; Pred. No. 3.3e-13;
Matches 93; Conservative 63; Mismatches 166; Indels 46; Gaps 12;

```
Qy          2 GPGEA--------LLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHL 53
              |||           : | | |::  |  :: |::| |    |   |  |    ||:|    |
Db         16 GPGPPPAPGGSGWVAAALCVVIVLTAAANSLLIVLICTQPAVRNT--SNFFLVSLFTSDL 73

Qy         54 LLAALDMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLR 113
              ::  : ||  :|  : ||   | | | :     |     | : |::  :| |::| :  |||
Db         74 MVGLVVMPPAMLNALYGRWVLARGLCLLWTAFDVMCCSASILNLCLISLDRYLLILSPLR 133

Qy        114 YAGRL-RPRYAGLLLGCAWGQSLA----FSGAALGCSWLGYSSAFASCSLRLPPEPERPR 168
              |  |:  ||   |:|| ||  |||    |    ||   || :           | | : |
Db        134 YKLRMTAPRALALILG-AW--SLAALASFLPLLLGWHELGKART---------PAPGQCR 181

Qy        169 FAA---FTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVT-------MKALALLA 218
                |   |     | | ||   :| |  ::   ||:  :: ::|       :: | :
Db        182 LLASLPFVLVASGVTFFLPSGAICFTYCRILLAARKQAVQVASLTTGTAGQALETLQVPR 241

Qy        219 DLHPSVR----QRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQ 274
                | :     :|   :  |:  :|:  :|| :  |  : : |: :  :|: |
Db        242 TPRPGMESADSRRLATKHSRKALKASLTLGILLGMFFVTWLPFFVANIAQAVCDCISPGL 301

Qy        275 WGILSKCLTYSKAVADPFTYSLLRRPFRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVH 334
              : :|:  | |  :  :|  | |  | |:: |     | |  :  | ||    | :   :
Db        302 FDVLT-WLGYCNSTMNPIIYPLFMRDFKRALG----RFLHASTVPRSTGQPCLPLHVDLS 356

Qy        335 QLLKRTPR 342
              |  :  |:
Db        357 QRCQTRPQ 364
```

RESULT 2
JC5520

serotonin receptor 6 - human
N;Alternate names: 5-hydroxytryptamine receptor 6 (5-HTR6)
C;Species: Homo sapiens (man)
C;Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 05-Nov-1999
C;Accession: JC5520
R;Kohen, R.; Metcalf, M.A.; Khan, N.; Druck, T.; Huebner, K.; Lachowicz, J.E.; Meltzer,
J. Neurochem. 66, 47-56, 1996
A;Title: Cloning, characterization, and chromosomal localization of a human 5-HT 6 serot
A;Reference number: JC5520; MUID:96102917; PMID:8522988
A;Accession: JC5520
A;Molecule type: mRNA
A;Residues: 1-440 <KOH>
A;Cross-references: GB:L41147; NID:g1162923; PIDN:AAA92622.1; PID:g1162924
A;Experimental source: brain
C;Comment: This protein shows high affinity for several therapeutically important antide
C;Genetics:
A;Gene: HTR6
A;Map position: 1p35-36
A;Introns: 238/3; 291/3
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transme
F;28-51/Domain: transmembrane #status predicted <TM1>
F;65-88/Domain: transmembrane #status predicted <TM2>
F;100-123/Domain: transmembrane #status predicted <TM3>
F;143-166/Domain: transmembrane #status predicted <TM4>
F;185-208/Domain: transmembrane #status predicted <TM5>
F;266-289/Domain: transmembrane #status predicted <TM6>
F;298-321/Domain: transmembrane #status predicted <TM7>
F;10/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.2%; Score 244.5; DB 2; Length 440;
Best Local Similarity 26.1%; Pred. No. 6.5e-13;
Matches 102; Conservative 64; Mismatches 164; Indels 61; Gaps 17;

```
Qy          2 GP-----GEALLAGLLVMVLAVALLSNALVL-LCCAYSAELRTRASGVLLVNLSLGHLLL 55
              ||     |   :|  | :|:|:   :|:|:: | |   |   |  |   ||:|    |::
Db         18 GPPSAPGGSGWVAAALCVVIALTAAANSLLIALICTQPALRNT--SNFFLVSLFTSDLMV 75

Qy         56 AALDMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYA 115
                : ||  :|  : ||   | | | :    |    | : |::  :| |::| :  ||||
Db         76 GLVVMPPAMLNALYGRWVLARGLCLLWTAFDVMCCSASILNLCLISLDRYLLILSPLRYK 135

Qy        116 GRLRP-RYAGLLLGCAWGQSLA----FSGAALGCSWLGYSSAFASCSLRLPPEPERPRFA 170
               |: | |   |:|| ||  |||    |    ||   ||::           || | : |
Db        136 LRMTPLRALALVLG-AW--SLAALASFLPLLLGWHELGHAR---------PPVPGQCRLL 183

Qy        171 A---FTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVT-------MKALALLADL 220
              |   |     : | ||   :| |  ::   ||:  :: ::|       : | :
Db        184 ASLPFVLVASGLTFFLPSGAICFTYCRILLAARKQAVQVASLTTGMASQASETLQVPRTP 243

Qy        221 HPSVR----QRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWG 276
              | |     :|   :  |:  :|:  :|| :  | : : |: : : :|       :
Db        244 RPGVESADSRRLATKHSRKALKASLTLGILLGMFFVTWLPFFVANIVQAVCDCISPGLFD 303

Qy        277 ILSKCLTYSKAVADPFTYSLLRRPFRQVLAGMVHRLL--KRTP--RPASTHDSSLDVA-- 330
              :|:  |  |  :  :|  | |  | |:: |    | |   | |  | ||   ||  :
Db        304 VLT-WLGYCNSTMNPIIYPLFMRDFKRALG----RFLPCPRCPRERQASLASPSLRTSHS 358

Qy        331 ----GMVHQLLKRTPRPASTHNGSVDTENDS 357
                  |:  | :   | |        |:::||
Db        359 GPRPGLSLQQVLPLPLPP-------DSDSDS 382
```

RESULT 3
JN0605
somatostatin receptor 4 - human
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000
C;Accession: JN0605; JN0762; A47457
R;Xu, Y.; Song, J.; Bruno, J.F.; Berelowitz, M.
Biochem. Biophys. Res. Commun. 193, 648-652, 1993
A;Title: Molecular cloning and sequencing of a human somatostatin receptor, hSSTR4.
A;Reference number: JN0605; MUID:93290656; PMID:8512564
A;Accession: JN0605
A;Molecule type: DNA
A;Residues: 1-388 <XUY>
A;Cross-references: GB:L14856; NID:g292499; PIDN:AAA36623.1; PID:g292500
R;Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.; Li
Biochem. Biophys. Res. Commun. 195, 844-852, 1993
A;Title: Cloning, functional expression and pharmacological characterization of a fourth
A;Reference number: JN0762; MUID:93384611; PMID:8373420
A;Accession: JN0762
A;Molecule type: DNA
A;Residues: 1-388 <YAM>
A;Cross-references: GB:D16826; NID:g693907; PIDN:BAA04106.1; PID:g693908
R;Rohrer, L.; Raulf, F.; Bruns, C.; Buettner, R.; Hofstaedter, F.; Schule, R.
Proc. Natl. Acad. Sci. U.S.A. 90, 4196-4200, 1993
A;Title: Cloning and characterization of a fourth human somatostatin receptor.
A;Reference number: A47457; MUID:93248256; PMID:8483934
A;Accession: A47457
A;Molecule type: DNA
A;Residues: 1-82,'T',84-364,'K',366-388 <ROH>
A;Cross-references: GB:L07833; NID:g307429; PIDN:AAA60565.1; PID:g307430
A;Note: sequence extracted from NCBI backbone (NCBIN:130856, NCBIP:130858)
C;Comment: This protein mediates the diverse actions of the tetradecaptide somatostatin.
C;Genetics:
A;Gene: GDB:SSTR4
A;Cross-references: GDB:202662; OMIM:182454
A;Map position: 20p11.2-20p11.2
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; phos
F;47-73/Domain: transmembrane #status predicted <TM1>
F;84-109/Domain: transmembrane #status predicted <TM2>
F;121-142/Domain: transmembrane #status predicted <TM3>
F;162-184/Domain: transmembrane #status predicted <TM4>
F;208-238/Domain: transmembrane #status predicted <TM5>
F;257-284/Domain: transmembrane #status predicted <TM6>
F;291-314/Domain: transmembrane #status predicted <TM7>
F;24/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;119-198/Disulfide bonds: #status predicted
F;161,253/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status pre
F;327/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 13.1%; Score 243.5; DB 2; Length 388;
Best Local Similarity 25.5%; Pred. No. 6.9e-13;
Matches 84; Conservative 55; Mismatches 124; Indels 67; Gaps 10;

```
Qy          2 GPGEALLAGLL------VMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLL 55
              |||:|  ||::       :|  | |: ||||:       |:::| |: : |:||::   |
Db         38 GPGDARAAGMVAIQCIYALVCLVGLVGNALVIFVILRYAKMKT-ATNIYLLNLAVADELF 96

Qy         56 AALDMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYA 115
                | :||          |      |: :  :|       :   :  || |:::||  ||| |
Db         97 -MLSVPFVASSAALRHWPFGSVLCRAVLSVDGLNMFTSVFCLTVLSVDRYVAVVHPLRAA 155

Qy        116 GRLRPRYAGLLLGCAWGQSL--------------AFSGAALGCS--WLGYSSAFASCSLR 159
               ||  | |:    |  ||               |  | |: |:  |
Db        156 TYRRPSVAKLINLGVWLASLLVTLPIAIFADTRPARGGQAVACNLQW------------- 202

Qy        160 LPPEPERPRFAAFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLAD 219
                |    | |      :||:||:  :  |             |  :    |:|:|| |
Db        203 ----PHPAWSAVFVVYTFLLGFLLPVLAIGL-----------CYLLIVGKMRAVALRAG 246

Qy        220 LHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILS 279
                      |::|   : || : : :  |::|: |: : :|  ||   :::|    :|
Db        247 WQ---------QRRRSEKKITRLVLMVVVVFVLCWMPFYVVQLLNLV-VTSLDATVNHVS 296

Qy        280 KCLTYSKAVADPFTYSLL----RRPFRQVL 305
              |:|: : |:|  |  |    || |::||
Db        297 LILSYANSCANPILYGFLSDNFRRSFQRVL 326
```

RESULT 4
JN0591
serotonin receptor 6 - rat

```
N;Alternate names: 5-hydroxytryptamine receptor 6 (5-HTR6)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
C;Accession: JN0591
R;Ruat, M.; Traiffort, E.; Arrang, J.M.; Tardivel-Lacombe, J.; Diaz, J.; Leurs, R.; Schw
Biochem. Biophys. Res. Commun. 193, 268-276, 1993
A;Title: A novel rat serotonin (5-HT6) receptor: molecular cloning, localization and sti
A;Reference number: JN0591; MUID:93277562; PMID:8389146
A;Accession: JN0591
A;Molecule type: DNA
A;Residues: 1-436 <RUA>
A;Cross-references: GB:S62043; NID:g385708; PIDN:AAB26908.1; PID:g385709
C;Genetics:
A;Introns: 238/3
C;Superfamily: vertebrate rhodopsin
C;Keywords: glycoprotein; neurotransmitter receptor; transmembrane protein
F;29-53/Domain: transmembrane #status predicted <TM1>
F;63-84/Domain: transmembrane #status predicted <TM2>
F;95-122/Domain: transmembrane #status predicted <TM3>
F;141-168/Domain: transmembrane #status predicted <TM4>
F;185-213/Domain: transmembrane #status predicted <TM5>
F;267-293/Domain: transmembrane #status predicted <TM6>
F;297-319/Domain: transmembrane #status predicted <TM7>
F;9/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             12.9%;  Score 239;  DB 2;  Length 436;
  Best Local Similarity   25.7%;  Pred. No. 1.9e-12;
  Matches   85;  Conservative   58;  Mismatches  146;  Indels   42;  Gaps   11;

Qy          2 GPGEA--------LLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHL 53
              |||          : | | |:::  |  :: |::| |    |   |  |   ||:|    |
Db         16 GPGPPPAPGGSGWVAAALCVVIVLTAAANSLLIVLICTQPALRNT--SNFFLVSLFTSDL 73

Qy         54 LLAALDMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLR 113
              ::  : ||  :|  : ||   | | | :    |    | : |::  :| |::| :  |||
Db         74 MVGLVVMPPAMLNALYGRWVLARGLCLLWTAFDVMCCSASILNLCLISLDRYLLILSPLR 133

Qy        114 YAGRL-RPRYAGLLLGCAWGQSLA----FSGAALGCSWLGYSSAFASCSLRLPPEPERPR 168
              |  |:  ||   |:|| ||  |||    |    ||   || :           | | : |
Db        134 YKLRMTAPRALALILG-AW--SLAALASFLPLLLGWHELGKART---------PAPGQCR 181

Qy        169 FAA---FTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVT-------MKALALLA 218
              |   |      | | ||   :| |  ::   ||:   :: ::|       :: | :
Db        182 LLASLPFVLVASGVTFFLPSGAICFTYCRILLAARKQAVQVASLTTGTAGQALETLQVPR 241

Qy        219 DLHPSVR----QRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQ 274
                 :     :|   :  |:  :|:  :|| :  | : : |: :  :|: |
Db        242 TPRPGMESADSRRLATKHSRKALKASLTLGILLGMFFVTWLPFFVANIAQAVCDCISPGL 301

Qy        275 WGILSKCLTYSKAVADPFTYSLLRRPFRQVL 305
              : :|:  | |  :  :|  | |  | |:: |
Db        302 FDVLT-WLGYCNSTMNPIIYPLFMRDFKRAL 331


RESULT 5
JC4120
histamine H2 receptor - guinea pig
C;Species: Cavia porcellus (guinea pig)
C;Date: 27-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 24-Nov-1999
C;Accession: JC4120
R;Traiffort, E.; Vizuete, M.L.; Tardivel-Lacombe, J.; Souil, E.; Schwartz, J.C.; Ruat, M
Biochem. Biophys. Res. Commun. 211, 570-577, 1995
A;Title: The guinea pig histamine H2 receptor: gene cloning, tissue expression and chrom
A;Reference number: JC4120; MUID:95314628; PMID:7794271
A;Accession: JC4120
A;Molecule type: DNA
A;Residues: 1-359 <TRA>
C;Comment: Histamine, a messenger molecule in cell-to-cell communication, affects its ta
ly defined receptor subtypes named H1, H2 and H3.
C;Superfamily: vertebrate rhodopsin
C;Keywords: glycoprotein; phosphoprotein; receptor; transmembrane protein
F;22-45/Domain: transmembrane #status predicted <TM1>
F;58-81/Domain: transmembrane #status predicted <TM2>
F;93-113/Domain: transmembrane #status predicted <TM3>
F;136-157/Domain: transmembrane #status predicted <TM4>
F;179-204/Domain: transmembrane #status predicted <TM5>
F;235-236/Domain: transmembrane #status predicted <TM6>
F;269-289/Domain: transmembrane #status predicted <TM7>
F;4,162,168/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;221,316/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicte
F;226,357/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicte

  Query Match             12.5%;  Score 232.5;  DB 2;  Length 359;
  Best Local Similarity   22.0%;  Pred. No. 5.4e-12;
  Matches   73;  Conservative   67;  Mismatches  139;  Indels   53;  Gaps    8;

Qy         11 LLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGVMRG 70
              :|:::: | :  | :| |    :  ||:  :   :|:|::  |||  | :||: : :
Db         24 ILIILILVTVAGNVVVCLAVGLNRRLRS-LTNCFIVSLAVTDLLLGLLVLPFSAIYQLSC 82

Qy         71 RTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLLLGCA 130
              :  :   | :   ||  | : : |::  :| |:: ||  ||||   : |    : |
Db         83 KWSFSKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPARVAISLVFI 142

Qy        131 WGQSLAFSGAALGCSWLGY------SSAFASCSLRLPPEPERPRFAAFTATLHAVGFVLP 184
              |  |:  |  ::   |         :     | :::           :       | | ||
Db        143 WVISITLSFLSIHLGWNSRNETSKDNDTIVKCKVQVN--------EVYGLVDGLVTFYLP 194

Qy        185 LAVLCLTSLQVHRVARRHCQRMDTV-TMKALALLADLHPSVRQRCLIQQKRRRHRATRKI 243
              | ::|:|  :: ::||   :|:: : : ||  :                   | |:||  :
Db        195 LLIMCITYFRIFKIAREQARRINHIGSWKAATI-----------------REHKATVTL 236

Qy        244 GIAIATFLICFAPYVMTRLAELVPFV--------TVNAQWGILSKCLTYSKAVADPFTYS 295
                :  |:||: ||          ||          ||  :  :   | |: :  :|  |:
Db        237 AAVMGAFIICWFPY-------FTVFVYRGLKGDDAVNEVFEDVVLWLGYANSALNPILYA 289

Qy        296 LLRRPFRQVLAGMVHRLLKRTPRPASTHDSSL 327
              | | ||      |:|       ::|::||
Db        290 ALNRDFRTA----YHQLFCCRLASHNSHETSL 317


RESULT 6
JN0763
somatostatin receptor 5 - human
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000
C;Accession: JN0763
R;Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.; Li
Biochem. Biophys. Res. Commun. 195, 844-852, 1993
A;Title: Cloning, functional expression and pharmacological characterization of a fourth
A;Reference number: JN0762; MUID:93384611; PMID:8373420
A;Accession: JN0763
A;Molecule type: DNA
A;Residues: 1-364 <YAM>
A;Cross-references: DDBJ:D16827; NID:g487683; PIDN:BAA04107.1; PID:g487684
C;Comment: This protein is a member of somatostatin receptor family.
C;Genetics:
A;Gene: GDB:SSTR5
A;Cross-references: GDB:138452; OMIM:182455
A;Map position: 16p13.3-16p13.3
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; lipoprotein; phosphoprotein; thiole
F;40-66/Domain: transmembrane #status predicted <TM1>
F;77-102/Domain: transmembrane #status predicted <TM2>
F;114-135/Domain: transmembrane #status predicted <TM3>
F;155-177/Domain: transmembrane #status predicted <TM4>
F;196-228/Domain: transmembrane #status predicted <TM5>
F;246-273/Domain: transmembrane #status predicted <TM6>
F;280-307/Domain: transmembrane #status predicted <TM7>
F;13,26,187/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;112-186/Disulfide bonds: #status predicted
F;242,325/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status pre
F;247/Binding site: phosphate (Thr) (covalent) (by cAMP-dependent kinase) #status predict
F;320/Binding site: palmitate (Cys) (covalent) #status predicted
```

```
  Query Match           12.5%;  Score 231.5;  DB 2;  Length 364;
  Best Local Similarity  25.0%;  Pred. No. 6.6e-12;
  Matches   92;  Conservative   51;  Mismatches  166;  Indels   59;  Gaps   11;

Qy          2 GPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMP 61
              |    |:  | ::| |  |  | ||:      |:::|  : : ::||::  :|   | :|
Db         37 GARAVLVPVLYLLVCAAGLGGNTLVIYVVLRFAKMKT-VTNIYILNLAVADVLY-MLGLP 94

Qy         62 FTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPR 121
              |          |  |  |::: ||      :   :  :| |::||| ||  |   |||
Db         95 FLATQNAASFWPFGPVLCRLVMTLDGVNQFTSVFCLTVMSVDRYLAVVHPLSSARWRRPR 154

Qy        122 YAGLLLGCAWGQSLAFS----------GAALGCSWLGYSSAFASCSLRLPPEPERPRFA 170
              | |    ||  ||  |           |    ||              |||     |
Db        155 VAKLASAAAWVLSLCMSLPLLVFADVQEGGTCNASW--------------PEPVGLWGA 199

Qy        171 AFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLI 230
              |      :||  || |:||             |  :  | ::|          ||  |:
Db        200 VFIIYTAVLGFFAPLLVICL-----------CYLLIVVKVRA--------AGVRVGCV- 238

Qy        231 QQKRRRHRATRKIGIAIATFLICFAPYV---MTRLAELVPFVTVNAQWGILSKCLTYSKA 287
               ::|   : || : : :  |  |: |:    :  ||  :|    :|        |:|: :
Db        239 -RRRSERKVTRMVLVVVLVFAGCWLPFFTVNIVNLAVALPQEPASAGLYFFVVILSYANS 297

Qy        288 VADPFTYSLLRRPFRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPA--S 345
              |:|  |  |   |||    :   |  |     |   |::      :  |  : || ||  :
Db        298 CANPVLYGFLSDNFRQSFQKV---LCLRKGSGAKDADATEPRPDRIRQQQEATP-PAHRA 353

Qy        346 THNGSVDT 353
                || : |
Db        354 AANGLMQT 361


RESULT 7
I57955
somatostatin receptor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999
C;Accession: I57955
R;Panetta, R.; Greenwood, M.T.; Warszynska, A.; Demchyshyn, L.L.; Day, R.; Niznik, H.B.;
Mol. Pharmacol. 45, 417-427, 1994
A;Title: Molecular cloning, functional characterization, and chromosomal localization of
A;Reference number: I57955; MUID:94195267; PMID:7908405
A;Accession: I57955
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-363 <RES>
A;Cross-references: GB:L14865; NID:g431094; PIDN:AAA20828.1; PID:g431095
C;Genetics:
A;Gene: GDB:SST
A;Cross-references: GDB:119604; OMIM:182450
A;Map position: 3q28-3q28
C;Superfamily: vertebrate rhodopsin

  Query Match           12.4%;  Score 231;  DB 2;  Length 363;
  Best Local Similarity  25.1%;  Pred. No. 7.2e-12;
  Matches   92;  Conservative   51;  Mismatches  166;  Indels   58;  Gaps   11;

Qy          2 GPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMP 61
              |    |:  | ::| |  |  | ||:      |:::|  : : ::||::  :|   | :|
Db         37 GARAVLVPVLYLLVCAAGLGGNTLVIYVVLRFAKMKT-VTNIYILNLAVADVLY-MLGLP 94

Qy         62 FTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPR 121
              |          |  |  |::: ||      :   :  :| |::||| ||  |   |||
Db         95 FLATQNAASFWPFGPVLCRLVMTLDGVNQFTSVFCLTVMSVDRYLAVVHPLSSARWRRPR 154

Qy        122 YAGLLLGCAWGQSLAFS----------GAALGCSWLGYSSAFASCSLRLPPEPERPRFA 170
              | |    ||  ||  |           |    ||              |||     |
Db        155 VAKLASAAAWVLSLCMSLPLLVFADVQEGGTCNASW--------------PEPVGLWGA 199

Qy        171 AFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLI 230
              |      :||  || |:||             |  :  | ::|          ||  |:
Db        200 VFIIYTAVLGFFAPLLVICL-----------CYLLIVVKVRA--------AGVRVGCV- 238

Qy        231 QQKRRRHRATRKIGIAIATFLICFAPYV---MTRLAELVPFVTVNAQWGILSKCLTYSKA 287
               ::|   : || : : :  |  |: |:    :  ||  :|    :|        |:|: :
Db        239 -RRRSERKVTRMVLVVVLVFAGCWLPFFTVNIVNLAVALPQEPASAGLYFFVVILSYANS 297

Qy        288 VADPFTYSLLRRPFRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRT-PRPAST 346
              |:|  |  |   |||    :   |  |     |   |::      :  |  : || ||  :
Db        298 CANPVLYGFLSDNFRQSFQKV---LCLRKGSGAKDADATEPRPDRIRQQQEATRPRTAAA 354

Qy        347 HNGSVDT 353
                || : |
Db        355 -NGLMQT 360


RESULT 8
QRHUBE
beta-3-adrenergic receptor, splice form 3 - human
N;Alternate names: beta-3-adrenergic receptor form B
N;Contains: beta-3-adrenergic receptor splice form 1
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1994 #sequence_revision 18-Aug-1995 #text_change 16-Jun-2000
C;Accession: S33751; S33753; S32803; S32826
R;Lelias, J.M.; Kaghad, M.; Rodriguez, M.; Chalon, P.; Bonnin, J.; Dupre, I.; Delpech, B.
FEBS Lett. 324, 127-130, 1993
A;Title: Molecular cloning of a human beta-3-adrenergic receptor cDNA.
A;Reference number: S33751; MUID:93285320; PMID:8389717
A;Accession: S33751
A;Molecule type: mRNA
A;Residues: 1-408 <LEL1>
A;Cross-references: EMBL:X70811; NID:g312396; PIDN:CAA50141.1; PID:g312397
A;Note: splice form 3
A;Accession: S33753
A;Molecule type: DNA
A;Residues: 392-408 <LEL2>
A;Cross-references: GB:X70812; NID:g312398; PIDN:CAA50143.1; PID:g1666376
R;Emorine, L.J.
submitted to the EMBL Data Library, March 1993
A;Reference number: S32803
A;Accession: S32803
A;Molecule type: DNA
A;Residues: 1-408 <EMO>
A;Cross-references: EMBL:X72861; NID:g298094; PIDN:CAA51383.1; PID:g298095
R;van Spronsen, A.; Nahmias, C.; Krief, S.; Briend-Sutren, M.M.; Strosberg, A.D.; Emorine
Eur. J. Biochem. 213, 1117-1124, 1993
A;Title: The promoter and intron/exon structure of the human and mouse beta3-adrenergic-
A;Reference number: S32826; MUID:93279311; PMID:8389293
A;Accession: S32826
A;Molecule type: DNA
A;Residues: 1-8;344-350;394-408 <SPR>
C;Genetics:
A;Gene: GDB:ADRB3
A;Cross-references: GDB:203869; OMIM:109691
A;Map position: 8p12-8p11.1
A;Introns: 402/2
C;Superfamily: vertebrate rhodopsin
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembrane
F;1-408/Product: beta-3-adrenergic receptor precursor splice form 3 #status predicted <M
F;1-402/Product: beta-3-adrenergic receptor precursor splice form 1 #status predicted <M
F;37-63/Domain: transmembrane #status predicted <TM1>
F;73-101/Domain: transmembrane #status predicted <TM2>
F;113-133/Domain: transmembrane #status predicted <TM3>
F;156-178/Domain: transmembrane #status predicted <TM4>
F;204-225/Domain: transmembrane #status predicted <TM5>
F;293-314/Domain: transmembrane #status predicted <TM6>
F;327-347/Domain: transmembrane #status predicted <TM7>
F;8,26/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match           12.4%;  Score 230.5;  DB 1;  Length 408;
  Best Local Similarity  24.9%;  Pred. No. 9e-12;
  Matches   93;  Conservative   61;  Mismatches  154;  Indels   65;  Gaps   14;

Qy          3 PGEALLAGLLVMVLAVALLS----NALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAAL 58
              | || ||| |   ||:|:|:    | ||::  |::  |:|  : | : :|:   |::  |
```

```
Db         34 PWEAALAGAL---LALAVLATVGGNLLVIVAIAWTPRLQT-MTNVFVTSLAAADLVMGLL 89

Qy         59 DMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRL 118

Db         90 VVPPAATLALTGHWPLGATGCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALV 149

Qy        119 RPRYAGLLLGCAWGQSLAFSGAALGCSW--LGYSSAFASCSLRLPPEPERPRFAAFTATL 176

Db        150 TKRCARTAVVLVWVVSAAVSFAPIMSQWWRVGADAEAQRCH-------SNPRCCAFASNM 202

Qy        177 ------HAVGFVLPLAVLCLTSLQVHRVARRHC-----------QRMDTVTMKALA--- 215

Db        203 PYVLLSSSVSFYLPLLVMLFVYARVFVVATRQLRLLRGELGRFPPEESPPAPSRSLAPAP 262

Qy        216 ----LLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRL------AEL 265

Db        263 VGTCAPPEGVPACGRRPARLLPLREHRALCTLGLIMGTFTLCWLPFFLANVLRALGGPSL 322

Qy        266 VP---FVTVNAQWGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTPRP 319

Db        323 VPGPAFLALN--W------LGYANSAFNPLIYCRSPDFRSAFRRLLCRCGRRL---PPEP 371

Qy        320 ASTHDSSLDVAGM 332

Db        372 CAAARPALFPSGV 384


RESULT 9
QRHUB3
beta-3-adrenergic receptor, splice form 2 - human
N;Alternate names: beta-3-adrenergic receptor form A
N;Contains: beta-3-adrenergic receptor splice form 1
C;Species: Homo sapiens (man)
C;Date: 22-Jan-1993 #sequence_revision 18-Aug-1995 #text_change 16-Jun-2000
C;Accession: A41348; S33752
R;Emorine, L.J.; Marullo, S.; Briend-Sutren, M.M.; Patey, G.; Tate, K.; Delavier-Klutchk
Science 245, 1118-1121, 1989
A;Title: Molecular characterization of the human beta-3-adrenergic receptor.
A;Reference number: A41348; MUID:89368947; PMID:2570461
A;Accession: A41348
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-402 <EMO>
A;Cross-references: GB:M29932; NID:g178895; PIDN:AAA35550.1; PID:g178896
A;Note: splice form 1
R;Lelias, J.M.; Kaghad, M.; Rodriguez, M.; Chalon, P.; Bonnin, J.; Dupre, I.; Delpech, B
FEBS Lett. 324, 127-130, 1993
A;Title: Molecular cloning of a human beta-3-adrenergic receptor cDNA.
A;Reference number: S33751; MUID:93285320; PMID:8389717
A;Accession: S33752
A;Molecule type: DNA
A;Residues: 392-414 <LEL>
A;Cross-references: EMBL:X70812; NID:g312398; PIDN:CAA50142.1; PID:g1666375
A;Note: splice form 2
C;Genetics:
A;Gene: GDB:ADRB3
A;Cross-references: GDB:203869; OMIM:109691
A;Map position: 8p12-8p11.1
A;Introns: 402/2
C;Superfamily: vertebrate rhodopsin
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembran
F;1-414/Product: beta-3-adrenergic receptor precursor splice form 2 #status predicted <M
F;1-402/Product: beta-3-adrenergic receptor precursor splice form 1 #status predicted <M
F;37-63/Domain: transmembrane #status predicted <TM1>
F;73-101/Domain: transmembrane #status predicted <TM2>
F;113-133/Domain: transmembrane #status predicted <TM3>
F;156-178/Domain: transmembrane #status predicted <TM4>
F;204-225/Domain: transmembrane #status predicted <TM5>
F;293-314/Domain: transmembrane #status predicted <TM6>
F;327-347/Domain: transmembrane #status predicted <TM7>
F;8,26/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             12.4%;  Score 230.5;  DB 1;  Length 414;
  Best Local Similarity   24.9%;  Pred. No. 9.2e-12;
  Matches   93;  Conservative   61;  Mismatches  154;  Indels   65;  Gaps   14;

Qy          3 PGEALLAGLLVMVLAVALLS----NALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAAL 58

Db         34 PWEAALAGAL---LALAVLATVGGNLLVIVAIAWTPRLQT-MTNVFVTSLAAADLVMGLL 89

Qy         59 DMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRL 118

Db         90 VVPPAATLALTGHWPLGATGCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALV 149

Qy        119 RPRYAGLLLGCAWGQSLAFSGAALGCSW--LGYSSAFASCSLRLPPEPERPRFAAFTATL 176

Db        150 TKRCARTAVVLVWVVSAAVSFAPIMSQWWRVGADAEAQRCH-------SNPRCCAFASNM 202

Qy        177 ------HAVGFVLPLAVLCLTSLQVHRVARRHC-----------QRMDTVTMKALA--- 215

Db        203 PYVLLSSSVSFYLPLLVMLFVYARVFVVATRQLRLLRGELGRFPPEESPPAPSRSLAPAP 262

Qy        216 ----LLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRL------AEL 265

Db        263 VGTCAPPEGVPACGRRPARLLPLREHRALCTLGLIMGTFTLCWLPFFLANVLRALGGPSL 322

Qy        266 VP---FVTVNAQWGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTPRP 319

Db        323 VPGPAFLALN--W------LGYANSAFNPLIYCRSPDFRSAFRRLLCRCGRRL---PPEP 371

Qy        320 ASTHDSSLDVAGM 332

Db        372 CAAARPALFPSGV 384


RESULT 10
G02953
beta-3-adrenergic receptor - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 13-Aug-1999
C;Accession: G02953
R;Lowe, A.L.; Walston, J.; Shuldiner, A.R.
submitted to the EMBL Data Library, July 1996
A;Reference number: H01989
A;Accession: G02953
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-418 <LOW>
A;Cross-references: EMBL:U63592; NID:g1458230; PIDN:AAB53939.1; PID:g1458232
C;Genetics:
A;Introns: 402/1
C;Superfamily: vertebrate rhodopsin

  Query Match             12.4%;  Score 229.5;  DB 2;  Length 418;
  Best Local Similarity   25.9%;  Pred. No. 1.1e-11;
  Matches   95;  Conservative   56;  Mismatches  149;  Indels   67;  Gaps   15;

Qy          3 PGEALLAGLLVMVLAVALLS----NALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAAL 58

Db         34 PWAAALAGAL---LALAVLATVGGNLLVIVAITRTPRLQT-MTNVFVTSLAAADLVMGLL 89

Qy         59 DMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRL 118

Db         90 VVPPAATLVLTGHWPLGATGCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALV 149

Qy        119 RPRYAGLLLGCAWGQSLAFSGAALGCSW--LGYSSAFASCSLRLPPEPERPRFAAFTATL 176

Db        150 TKRRARAAVVLVWVVSAAVSFAPIMSQWWRVGADAEAQRCH-------SNPRCCAFASNM 202

Qy        177 ------HAVGFVLPLAVLCLTSLQVHRVARRHCQRMD------TVTMKALALLADLHPSV 224

Db        203 PYVLLSSSVSFYLPLLVMLFVYARVFVVATRQLRLLRWELGRFPPEESSPALSRSLAPAP 262

Qy        225 RQRCLIQQ------------KRRRHRATRKIGIAIATFLICFAPYVMTRL------AEL 265

Db        263 AGTCAPPEGVPACCRRPARLLPLREHRALCTLGLIMGTFTLCWLPFFLANVLRALGGPSL 322
```

```
Qy        266 VP---FVTVNAQWGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTP-- 317
              ||   |: :|  |      | |: :  :|  |      |  ||::|     | | | |
Db        323 VPDPAFLALN--W------LGYANSAFNPLIYCRSPDFRSAFRRLLCHCGGR-LPREPCA 373

Qy        318 --RPAST 322
                 ||||:
Db        374 ADRPASS 380


RESULT 11
I57940
somatostatin receptor 5 - rat
N;Alternate names: somatotropin release-inhibiting factor subtype 28 receptor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 24-Nov-1999
C;Accession: I57940; I57949; S39244
R;O'Carroll, A.M.; Lolait, S.J.; Konig, M.; Mahan, L.C.
Mol. Pharmacol. 42, 939-946, 1992
A;Title: Molecular cloning and expression of a pituitary somatostatin receptor with pref
A;Reference number: I57940; MUID:93125499; PMID:1362243
A;Accession: I57940
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-363 <OCA1>
A;Cross-references: GB:L04535; NID:g409238; PIDN:AAA17029.1; PID:g409239
R;O'Carroll, A.M.; Lolait, S.J.; Konig, M.; Mahan, L.C.
Mol. Pharmacol. 44, 1278, 1993
A;Title: Molecular cloning and expression of a pituitary somatostatin receptor with pref
A;Reference number: I57949; MUID:94088493; PMID:8264565
A;Accession: I57949
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 341-363 <OCA2>
A;Cross-references: GB:S67370; NID:g455947; PIDN:AAB29371.1; PID:g455948
A;Experimental source: pituitary
R;Penetta, R.; Greenwood, M.; Patel, Y.C.
submitted to the EMBL Data Library, August 1993
A;Description: Correction of the nucleotide and amino acid sequence of the rat somatosta
A;Reference number: S39244
A;Accession: S39244
A;Molecule type: mRNA
A;Residues: 309-363 <PEN>
A;Cross-references: EMBL:X74828; NID:g433911; PIDN:CAA52825.1; PID:g433912
C;Genetics:
A;Gene: SSTR5
C;Superfamily: vertebrate rhodopsin

  Query Match            12.3%;  Score 228.5;  DB 2;  Length 363;
  Best Local Similarity  25.6%;  Pred. No. 1.2e-11;
  Matches   95;  Conservative   55;  Mismatches  160;  Indels   61;  Gaps   14;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              ||     |:  | ::|  | |  | ||:       |:::|  : | ::||::  :|   | :
Db         33 MGARAVLVPVLYLLVCTVGLSGNTLVIYVVLRHAKMKT-VTNVYILNLAVADVLF-MLGL 90

Qy         61 PF-TLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLR 119
              ||       :    |      |:::  ||       :   :  :| |::|||  ||| |   |
Db         91 PFLATQNAVVSYWPFGSFLCRLVMTLDGINQFTSIFCLMVMSVDRYLAVVHPLRSARWRR 150

Qy        120 PRYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAV 179
              || | :      |  ||  |   |    :      : :|:|   |||     |||      :
Db        151 PRVAKMASAAVWVFSLLMSLPLL--VFADVQEGWGTCNLSW-PEPVGLWGAAFITYTSVL 207

Qy        180 GFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRA 239
              ||  || |:||           |  :  | :||  :      | |:|       |   :
Db        208 GFFGPLLVICL------------CYLLIVVKVKAAGMRVG---SSRRR------RSEPKV 246

Qy        240 TRKIGIAIATFLICFAPYVMTRLAEL---VPFVTVNAQWGILSKCLTYSKAVADPFTYSL 296
              || : : :  |: |: |: :  :  |   :|     :|       |:|: : |:|  |
Db        247 TRMVVVVVLVFVGCWLPFFIVNIVNLAFTLPEEPTSAGLYFFVVVLSYANSCANPLLYGF 306

Qy        297 L----RRPFRQVLAGMVHRLLKR----------TPRPASTHDSSLDVAGMVHQLLKRTPR 342
              |    |: ||:||        |:|            |||         | :|     |   |
Db        307 LSDNFRQSFRKVLC------LRRGYGMEDADAIEPRP--------DKSGRPQATL---PT 349

Qy        343 PASTHNGSVDT 353
               :   || : |
Db        350 RSCEANGLMQT 360


RESULT 12
A41795
somatostatin receptor 1 - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 24-Nov-1999
C;Accession: A41795
R;Yamada, Y.; Post, S.R.; Wang, K.; Tager, H.S.; Bell, G.I.; Seino, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 251-255, 1992
A;Title: Cloning and functional characterization of a family of human and mouse somatosta
A;Reference number: A41795; MUID:92108031; PMID:1346068
A;Accession: A41795
A;Molecule type: DNA
A;Residues: 1-391 <YAM>
A;Cross-references: GB:M81829; NID:g307433; PIDN:AAA58247.1; PID:g307434
A;Note: sequence extracted from NCBI backbone (NCBIN:74767, NCBIP:74768)
C;Genetics:
A;Gene: GDB:SSTR1
A;Cross-references: GDB:134185; OMIM:182451
A;Map position: 14q13-14q13
A;Introns: #status absent
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; phos
F;58-84/Domain: transmembrane #status predicted <TM1>
F;95-120/Domain: transmembrane #status predicted <TM2>
F;132-153/Domain: transmembrane #status predicted <TM3>
F;173-195/Domain: transmembrane #status predicted <TM4>
F;220-250/Domain: transmembrane #status predicted <TM5>
F;269-296/Domain: transmembrane #status predicted <TM6>
F;302-326/Domain: transmembrane #status predicted <TM7>
F;4,44,48,381/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;130-208/Disulfide bonds: #status predicted
F;172/Binding site: phosphate (Thr) (covalent) (by cAMP-dependent kinase) #status predict
F;265/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predict
F;339/Binding site: palmitate (Cys) (covalent) #status predicted

  Query Match            12.3%;  Score 228.5;  DB 2;  Length 391;
  Best Local Similarity  24.5%;  Pred. No. 1.3e-11;
  Matches   78;  Conservative   61;  Mismatches  131;  Indels   49;  Gaps   11;

Qy          2 GPGEALLAGLLVMVLA-VALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              | | |:|   :  |:  | |  |::|:       |:::| |: : ::||::   ||  | :
Db         54 GQGSAILISFIYSVVCLVGLCGNSMVIYVILRYAKMKT-ATNIYILNLAIADELL-MLSV 111

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              || :   :    |      |:::  :|       :   :  || |:::||  |:: |   ||
Db        112 PFLVTSTLLRHWPFGALLCRLVLSVDAVNMFTSIYCLTVLSVDRYVAVVHPIKAARYRRP 171

Qy        121 RYAGLLLGCAWGQSL-------AFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFT 173
              | ::    |  ||         ||  |        |    :|:: | ||| :     |
Db        172 TVAKVVNLGVWVLSLLVILPIVVFSRTAAN------SDGTVACNM-LMPEPAQRWLVGFV 224

Qy        174 ATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQK 233
              :||:||:  :||          |  :   |: :|| |                 |:|
Db        225 LYTFLMGFLLPVGAICL------------CYVLIIAKMRMVALKAGWQ---------QRK 263

Qy        234 RRRHRATRKIGIAIATFLICFAPYVMTRLAELV---PFVTVNAQWGILSKCLTYSKAVAD 290
              |   : |  : : :  |:||: |: : :|  :       ||:     ||  | |: : |:
Db        264 RSERKITLMVMMVVMVFVICWMPFYVVQLVNVFAEQDDATVSQ----LSVILGYANSCAN 319

Qy        291 PFTYSLL----RRPFRQVL 305
              |  |  |    :| |:::|
Db        320 PILYGFLSDNFKRSFQRIL 338


RESULT 13
```

```
C41795
somatostatin receptor 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1993 #sequence_revision 31-Dec-1993 #text_change 24-Nov-1999
C;Accession: C41795
R;Yamada, Y.; Post, S.R.; Wang, K.; Tager, H.S.; Bell, G.I.; Seino, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 251-255, 1992
A;Title: Cloning and functional characterization of a family of human and mouse somatost
A;Reference number: A41795; MUID:92108031; PMID:1346068
A;Accession: C41795
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-391 <YAM>
A;Cross-references: GB:M81831; NID:g201058; PIDN:AAA58255.1; PID:g201059
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein

  Query Match             12.3%;  Score 228.5;  DB 2;  Length 391;
  Best Local Similarity   24.5%;  Pred. No. 1.3e-11;
  Matches   78;  Conservative   61;  Mismatches  131;  Indels   49;  Gaps   11;

Qy          2 GPGEALLAGLLVMVLA-VALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              | | |:|   :  |:  | |  |::|:      |:::| |: : ::||::   ||  | :
Db         54 GQGSAILISFIYSVVCLVGLCGNSMVIYVILRYAKMKT-ATNIYILNLAIADELL-MLSV 111

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              || :   :    |     |:::  :|       :   :  || |:::||  |:: |   ||
Db        112 PFLVTSTLLRHWPFGALLCRLVLSVDAVNMFTSIYCLTVLSVDRYVAVVHPIKAARYRRP 171

Qy        121 RYAGLLLGCAWGQSL-------AFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFT 173
                | ::    |  ||       ||  |       |    :|:: | ||| :     |
Db        172 TVAKVVNLGVWVLSLLVILPIVVFSRTAAN------SDGTVACNM-LMPEPAQRWLVGFV 224

Qy        174 ATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQK 233
                :||:||:  :||       |  :    |: :|| |           |:|
Db        225 LYTFLMGFLLPVGAICL-----------CYVLIIAKMRMVALKAGWQ---------QRK 263

Qy        234 RRRHRATRKIGIAIATFLICFAPYVMTRLAELV---PFVTVNAQWGILSKCLTYSKAVAD 290
              |   : |  : : :  |:||: |: : :|  :        ||:    ||  | |: : |:
Db        264 RSERKITLMVMMVVMVFVICWMPFYVVQLVNVFAEQDDATVSQ----LSVILGYANSCAN 319

Qy        291 PFTYSLL----RRPFRQVL 305
              |  |  |    :| |:::|
Db        320 PILYGFLSDNFKRSFQRIL 338


RESULT 14
A39297
somatostatin receptor - rat
N;Alternate names: probable G-protein-coupled receptor; SRIF receptor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 24-Nov-1999
C;Accession: A39297; A45102; S20088
R;Meyerhof, W.; Paust, H.J.; Schoenrock, C.; Richter, D.
DNA Cell Biol. 10, 689-694, 1991
A;Title: Cloning of a cDNA encoding a novel putative G-protein-coupled receptor expresse
A;Reference number: A39297; MUID:92096119; PMID:1661599
A;Accession: A39297
A;Molecule type: mRNA
A;Residues: 1-391 <MEY>
A;Cross-references: GB:X62314; GB:X61630; NID:g56309; PIDN:CAA44193.1; PID:g56310
A;Experimental source: brain
A;Note: it is uncertain whether Met-1 is the initiator or whether translation is initiat
R;Li, X.J.; Forte, M.; North, R.A.; Ross, C.A.; Snyder, S.H.
J. Biol. Chem. 267, 21307-21312, 1992
A;Title: Cloning and expression of a rat somatostatin receptor enriched in brain.
A;Reference number: A45102; MUID:93016064; PMID:1400442
A;Accession: A45102
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-391 <LI1>
A;Experimental source: brain
A;Note: sequence extracted from NCBI backbone (NCBIP:116692)
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; receptor; transmembrane protein

  Query Match             12.3%;  Score 228.5;  DB 2;  Length 391;
  Best Local Similarity   24.5%;  Pred. No. 1.3e-11;
  Matches   78;  Conservative   61;  Mismatches  131;  Indels   49;  Gaps   11;

Qy          2 GPGEALLAGLLVMVLA-VALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              | | |:|   :  |:  | |  |::|:      |:::| |: : ::||::   ||  | :
Db         54 GQGSAILISFIYSVVCLVGLCGNSMVIYVILRYAKMKT-ATNIYILNLAIADELL-MLSV 111

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              || :   :    |     |:::  :|       :   :  || |:::||  |:: |   ||
Db        112 PFLVTSTLLRHWPFGALLCRLVLSVDAVNMFTSIYCLTVLSVDRYVAVVHPIKAARYRRP 171

Qy        121 RYAGLLLGCAWGQSL-------AFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFT 173
                | ::    |  ||       ||  |       |    :|:: | ||| :     |
Db        172 TVAKVVNLGVWVLSLLVILPIVVFSRTAAN------SDGTVACNM-LMPEPAQRWLVGFV 224

Qy        174 ATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQK 233
                :||:||:  :||       |  :    |: :|| |           |:|
Db        225 LYTFLMGFLLPVGAICL-----------CYVLIIAKMRMVALKAGWQ---------QRK 263

Qy        234 RRRHRATRKIGIAIATFLICFAPYVMTRLAELV---PFVTVNAQWGILSKCLTYSKAVAD 290
              |   : |  : : :  |:||: |: : :|  :        ||:    ||  | |: : |:
Db        264 RSERKITLMVMMVVMVFVICWMPFYVVQLVNVFAEQDDATVSQ----LSVILGYANSCAN 319

Qy        291 PFTYSLL----RRPFRQVL 305
              |  |  |    :| |:::|
Db        320 PILYGFLSDNFKRSFQRIL 338


RESULT 15
DYHUD5
dopamine receptor D5 - human
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 22-Jun-1999
C;Accession: S15080; A41232
R;Sunahara, R.K.; Guan, H.C.; O'Dowd, B.F.; Seeman, P.; Laurier, L.G.; Ng, G.; George, S.
Nature 350, 614-619, 1991
A;Title: Cloning of the gene for a human dopamine D(5) receptor with higher affinity for
A;Reference number: S15080; MUID:91204055; PMID:1826762
A;Accession: S15080
A;Molecule type: DNA
A;Residues: 1-477 <SUN>
A;Cross-references: EMBL:X58454; NID:g32048; PIDN:CAA41360.1; PID:g32049
R;Grandy, D.K.; Zhang, Y.; Bouvier, C.; Zhou, Q.Y.; Johnson, R.A.; Allen, L.; Buck, K.; F
Proc. Natl. Acad. Sci. U.S.A. 88, 9175-9179, 1991
A;Title: Multiple human D-5 dopamine receptor genes: a functional receptor and two pseudo
A;Reference number: A41232; MUID:92021013; PMID:1833775
A;Accession: A41232
A;Molecule type: DNA
A;Residues: 1-477 <GRA>
A;Cross-references: GB:M67439; NID:g181830; PIDN:AAA52329.1; PID:g181831
C;Genetics:
A;Gene: GDB:DRD5; DRD1L2
A;Cross-references: GDB:127548; OMIM:126453
A;Map position: 4p15.3-4p15.1
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; lipoprotein; neurotransmitter recep
F;41-66/Domain: transmembrane #status predicted <TM1>
F;78-104/Domain: transmembrane #status predicted <TM2>
F;115-136/Domain: transmembrane #status predicted <TM3>
F;157-180/Domain: transmembrane #status predicted <TM4>
F;225-245/Domain: transmembrane #status predicted <TM5>
F;246-297/Domain: intracellular #status predicted <INT>
F;298-319/Domain: transmembrane #status predicted <TM6>
F;341-361/Domain: transmembrane #status predicted <TM7>
F;7,199/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;113-217/Disulfide bonds: #status predicted
F;153,292/Binding site: phosphate (Thr) (covalent) #status predicted
F;260,271,283/Binding site: phosphate (Ser) (covalent) #status predicted
F;375/Binding site: palmitate (Cys) (covalent) #status predicted
```

```
  Query Match            12.3%;  Score 228.5;  DB 1;  Length 477;
  Best Local Similarity  21.6%;  Pred. No. 1.6e-11;
  Matches   90;  Conservative   82;  Mismatches  162;  Indels   83;  Gaps   17;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              :|| : : | || :::   || | ||       |  ||   : | :|:|::  | :| | |
Db         36 LGPSQVVTACLLTLLIIWTLLGNVLVCAAIVRSRHLRANMTNVFIVSLAVSDLFVALLVM 95

Qy         61 PFTLLGVMRGRTPSAPGA-CQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLR 119
              |:  : : | |   || | |    |   :: : |::  :| |:: |:  | ||  ::
Db         96 PWKAVAEVAGYWPF--GAFCDVWVAFDIMCSTASILNLCVISVDRYWAISRPFRYKRKMT 153

Qy        120 PRYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPP------------EPE- 165
               | | :::| ||  |:  |   :  :|     | :   | ||             ||:
Db        154 QRMALVMVGLAWTLSILISFIPVQLNW-HRDQAASWGGLDLPNNLANWTPWEEDFWEPDV 212

Qy        166 --------RPRFAAFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKA---- 213
                     |  | :::|  : | :|:|:: :|  :::|:|:   :|: ::   |
Db        213 NAENCDSSLNRTYAISSSL--ISFYIPVAIMIVTYTRIYRIAQVQIRRISSLERAAEHAQ 270

Qy        214 -LALLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVN 272
                |   |    |  |   ::  :  : : : :  |: |: |: :  |  :||| : :
Db        271 SCRSSAACAPDTSLRASI---KKETKVLKTLSVIMGVFVCCWLPFFI--LNCMVPFCSGH 325

Qy        273 AQ-----------------WGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHR 311
              :                 |      :: :  :|  |:     :: | |:| |  |
Db        326 PEGPPAGFPCVSETTFDVFVW------FGWANSSLNPVIYAFNADFQKVFAQLL-GCSH- 377

Qy        312 LLKRTPRPASTHDSSL-----DV-------AGMVHQLLKRTPRPASTHNGSVDTEND 356
               |||      : |     |:       | :| :    |   :  |  || : :
Db        378 FCSRTPVETVNISNELISYNQDIVFHKEIAAAYIHMM----PNAVTPGNREVDNDEE 430


Search completed: May 18, 2004, 12:08:40
Job time : 21 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 18, 2004, 12:01:36 ; Search time 18 Seconds
(without alignments)
1050.081 Million cell updates/sec

Title: US-10-049-569-2
Perfect score: 1858
Sequence: 1 MGPGEALLAGLLVMVLAVAL.....ASTHNGSVDTENDSCLQQTH 363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1843	99.2	363	1	GP78_HUMAN	Q96p69 homo sapien
2	848	45.6	337	1	GP26_HUMAN	Q8ndv2 homo sapien
3	823	44.3	337	1	GP26_RAT	Q9qxi3 rattus norv
4	822	44.2	337	1	GP26_MOUSE	Q8bza7 mus musculu
5	245.5	13.2	440	1	5H6_MOUSE	Q9r1c8 mus musculu
6	244.5	13.2	440	1	5H6_HUMAN	P50406 homo sapien
7	243.5	13.1	388	1	SSR4_HUMAN	P31391 homo sapien
8	239	12.9	436	1	5H6_RAT	P31388 rattus norv
9	232.5	12.5	359	1	HH2R_CAVPO	P47747 cavia porce
10	232	12.5	388	1	5H4_MOUSE	P97288 mus musculu
11	231.5	12.5	364	1	SSR5_HUMAN	P35346 homo sapien
12	230.5	12.4	408	1	B3AR_HUMAN	P13945 homo sapien
13	229.5	12.4	418	1	B3AR_MACMU	Q28524 macaca mula
14	229.5	12.4	425	1	OX1R_HUMAN	O43613 homo sapien
15	229	12.3	362	1	SSR5_MOUSE	O08858 mus musculu
16	228.5	12.3	363	1	SSR5_RAT	P30938 rattus norv
17	228.5	12.3	391	1	SSR1_HUMAN	P30872 homo sapien
18	228.5	12.3	391	1	SSR1_MOUSE	P30873 mus musculu
19	228.5	12.3	391	1	SSR1_RAT	P28646 rattus norv
20	228.5	12.3	477	1	DBDR_HUMAN	P21918 homo sapien
21	228	12.3	387	1	GALS_HUMAN	O43603 homo sapien
22	228	12.3	475	1	DBDR_RAT	P25115 rattus norv
23	224.5	12.1	358	1	HH2R_MOUSE	P97292 mus musculu
24	224	12.1	372	1	GALS_RAT	O08726 rattus norv
25	224	12.1	457	1	DBDR_XENLA	P42290 xenopus lae
26	224	12.1	514	1	A1AB_MOUSE	P97717 mus musculu
27	223.5	12.0	359	1	HH2R_CANFA	P17124 canis famil
28	223	12.0	369	1	SSR2_HUMAN	P30874 homo sapien
29	223	12.0	388	1	5H4_CAVPO	O70528 cavia porce
30	223	12.0	467	1	B1AR_BOVIN	Q9tt96 bos taurus
31	223	12.0	479	1	5H7_HUMAN	P34969 homo sapien
32	222.5	12.0	384	1	SSR4_RAT	P30937 rattus norv
33	222.5	12.0	561	1	A1AD_RAT	P23944 rattus norv
34	222	11.9	369	1	SSR2_RAT	P30680 rattus norv
35	222	11.9	388	1	5H4_HUMAN	Q13639 homo sapien
36	222	11.9	448	1	5H7_RAT	P32305 rattus norv
37	222	11.9	459	1	D1DR_FUGRU	P53452 fugu rubrip
38	221.5	11.9	371	1	GALS_MOUSE	O88854 mus musculu
39	221	11.9	369	1	SSR2_PIG	P34994 sus scrofa
40	220	11.8	359	1	HH2R_HUMAN	P25021 homo sapien
41	220	11.8	359	1	HH2R_PANTR	P60021 pan troglod
42	220	11.8	370	1	5H5B_MOUSE	P31387 mus musculu
43	220	11.8	470	1	A1AA_ORYLA	Q91175 oryzias lat
44	220	11.8	562	1	A1AD_MOUSE	P97714 mus musculu
45	219	11.8	405	1	B3AR_CANFA	O02662 canis famil

ALIGNMENTS

RESULT 1
GP78_HUMAN
ID GP78_HUMAN STANDARD; PRT; 363 AA.
AC Q96P69; Q8NGV3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable G protein-coupled receptor GPR78.
GN GPR78.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21458557; PubMed=11574155;
RA Lee D.K., Nguyen T., Lynch K.R., Cheng R., Vanti W.B., Arkhitko O.,
RA Lewis T., Evans J.F., George S.R., O'Dowd B.F.;
RT "Discovery and mapping of ten novel G protein-coupled receptor
RT genes.";
RL Gene 275:83-91(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA Tsutsumi S., Aburatani H., Asai K., Akiyama Y.;
RT "Genome-wide discovery and analysis of human seven transmembrane helix
RT receptor genes.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP INDUCTION.
RX MEDLINE=20175424; PubMed=10708769;
RA van Laar T., Schouten T., Hoogervorst E., van Eck M., van der Eb A.J.,
RA Terleth C.;
RT "The novel MMS-inducible gene Mif1/KIAA0025 is a target of the
RT unfolded protein response pathway.";
RL FEBS Lett. 469:123-131(2000).
CC -!- FUNCTION: Orphan receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in pituitary and placenta. No
CC expression detected in brain, skeletal muscle, lung, heart, liver,
CC pancreas, or kidney.
CC -!- INDUCTION: Expression increased in fibroblasts or HeLa cells
CC following UV-A irradiation, exposure to DNA-alkylating agents, or
CC endoplasmic reticulum (ER) stress caused by osmotic shock or the
CC glycosylation inhibitor tunicamycin.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC ----------

CC ----------
DR EMBL; AF411107; AAL26479.1; -.

```
DR   EMBL; AB065673; BAC05898.1; -.
DR   Genew; HGNC:4528; GPR78.
DR   MIM; 606921; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane.
FT   DOMAIN        1      7       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      8     28       1 (POTENTIAL).
FT   DOMAIN       29     47       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     48     68       2 (POTENTIAL).
FT   DOMAIN       69     80       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     81    101       3 (POTENTIAL).
FT   DOMAIN      102    122       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    123    143       4 (POTENTIAL).
FT   DOMAIN      144    168       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    169    189       5 (POTENTIAL).
FT   DOMAIN      190    242       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    243    263       6 (POTENTIAL).
FT   DOMAIN      264    277       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    278    297       7 (POTENTIAL).
FT   DOMAIN      298    363       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     79    156       BY SIMILARITY.
FT   CONFLICT    216    216       V -> L (IN REF. 2).
FT   CONFLICT    226    227       HG -> QR (IN REF. 2).
SQ   SEQUENCE   363 AA;  39228 MW;  00E6D9FFE768E0ED CRC64;

  Query Match             99.2%;  Score 1843;  DB 1;  Length 363;
  Best Local Similarity   99.2%;  Pred. No. 4.4e-130;
  Matches  360;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |||||||||||||||||||||||||||||||:|||||||| ||||||||||||
Db        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALAVLADLHPSVRHGCLIQQKRRRHRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQ 360

Qy        361 QTH 363
              |||
Db        361 QTH 363


RESULT 2
GP26_HUMAN
ID   GP26_HUMAN     STANDARD;      PRT;   337 AA.
AC   Q8NDV2;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Probable G protein-coupled receptor GPR26.
GN   GPR26.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Jones G., Boulay J.L., Maier D., Merlo A.;
RT   "Sequence of the human homologue of the rat orphan G protein-coupled
RT   receptor GPR26 mRNA.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Boulay J.L., Labuhn M., Jones G., Maier D., Merlo A.;
RT   "The 10q25.3-26.1 gene encoding the orphan G protein-coupled receptor
RT   GPR26 is epigenetically silenced in human gliomas.";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Orphan receptor.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ505757; CAD44281.1; -.
DR   Genew; HGNC:4481; GPR26.
DR   MIM; 604847; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane.
FT   DOMAIN        1     10       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     11     31       1 (POTENTIAL).
FT   DOMAIN       32     47       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     48     68       2 (POTENTIAL).
FT   DOMAIN       69     81       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     82    102       3 (POTENTIAL).
FT   DOMAIN      103    123       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    124    144       4 (POTENTIAL).
FT   DOMAIN      145    168       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    169    189       5 (POTENTIAL).
FT   DOMAIN      190    245       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    246    266       6 (POTENTIAL).
FT   DOMAIN      267    276       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    277    297       7 (POTENTIAL).
FT   DOMAIN      298    337       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     79    156       BY SIMILARITY.
SQ   SEQUENCE   337 AA;  37603 MW;  4E04D371505F6D01 CRC64;

  Query Match             45.6%;  Score 848;  DB 1;  Length 337;
  Best Local Similarity   52.0%;  Pred. No. 3.8e-56;
  Matches  170;  Conservative   52;  Mismatches  101;  Indels    4;  Gaps    1;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              |   :| ||||||  : |:|||||||||  :||::| :|  :  :||: |:||   ::|
Db          1 MNSWDAGLAGLLVGTMGVSLLSNALVLLCLLHSADIRRQAPALFTLNLTCGNLLCTVVNM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              | || ||:  | |:    |::  |||||||:|: ||:|||| |:|:|| ||| |  ::|
Db         61 PLTLAGVVAQRQPAGDRLCRLAAFLDTFLAANSMLSMAALSIDRWVAVVFPLSYRAKMRL 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              | | |::   |  :| |  |||  ||||:   :|||:|      || ||| ||   ||:
Db        121 RDAALMVAYTWLHALTFPAAALALSWLGFHQLYASCTLCSRRPDERLRFAVFTGAFHALS 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |:|   ||| | |:| :||| ||:|:| :||: | || ||||||||:||| :||||| |||
```

```
Db        181 FLLSFVVLCCTYLKVLKVARFHCKRIDVITMQTLVLLVDLHPSVRERCLEEQKRRRQRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              :||    | |||:||||||:||| ||    | : : ||:||||| |||| :||| |||||
Db        241 KKISTFIGTFLVCFAPYVITRLVELFSTVPIGSHWGVLSKCLAYSKAASDPFVYSLLRHQ 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSL 327
              :|:    :::||| |      | | | |
Db        301 YRKSCKEILNRLLHR----RSIHSSGL 323


RESULT 3
GP26_RAT
ID   GP26_RAT       STANDARD;      PRT;   337 AA.
AC   Q9QXI3;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Probable G protein-coupled receptor GPR26.
GN   GPR26.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20149852; PubMed=10684976;
RA   Lee D.K., Lynch K.R., Nguyen T., Im D.-S., Cheng R., Saldivia V.R.,
RA   Liu Y., Liu I.S.C., Heng H.H.Q., Seeman P., George S.R., O'Dowd B.F.,
RA   Marchese A.;
RT   "Cloning and characterization of additional members of the G protein-
RT   coupled receptor family.";
RL   Biochim. Biophys. Acta 1490:311-323(2000).
CC   -!- FUNCTION: Orphan receptor.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Detected in extracts of several brain regions
CC       including striatum, pons, cerebellum and cortex. Not detected in
CC       numerous peripheral tissue extracts, except in testis. In the
CC       brain, detected in cortical structures including the anterior
CC       cingulate area, posterior cingulate and the frontoparietal,
CC       somatosensory and piriform cortices. Prominent also in the
CC       olfactory tubercle, the islands of Calleja, ventromedial and
CC       posterior nuclei of the hypothalamus, the medial septal nucleus,
CC       nucleus of the diagonal band and the ventral tegmental area.
CC       Localized also to hippocampal structures, with signals strongest
CC       over the CA2 and CA3 regions of Ammon's horn and less so over the
CC       dentate gyrus. Expressed in the caudate putamen only in its most
CC       caudal portion, with a decreasing gradient of signal from the
CC       dorsal to ventral aspect. Strong expression associated with a
CC       single pontine structure, the inferior olivary nucleus.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF208288; AAF21012.1; -.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR000169; SHprot_acsite.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane.
FT   DOMAIN        1     10       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     11     31       1 (POTENTIAL).
FT   DOMAIN       32     47       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     48     68       2 (POTENTIAL).
FT   DOMAIN       69     81       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     82    102       3 (POTENTIAL).
FT   DOMAIN      103    123       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    124    144       4 (POTENTIAL).
FT   DOMAIN      145    168       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    169    189       5 (POTENTIAL).
FT   DOMAIN      190    245       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    246    266       6 (POTENTIAL).
FT   DOMAIN      267    276       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    277    297       7 (POTENTIAL).
FT   DOMAIN      298    337       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     79    156       BY SIMILARITY.
SQ   SEQUENCE   337 AA;  37804 MW;  CD7E2381F085F2DD CRC64;

  Query Match           44.3%;  Score 823;  DB 1;  Length 337;
  Best Local Similarity   49.1%;  Pred. No. 2.7e-54;
  Matches  164;  Conservative   58;  Mismatches  106;  Indels    6;  Gaps    2;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              |   :| |||||| : |:|||| |||||  :||::| :|  :  :||: |:||   ::|
Db          1 MNSWDAGLAGLLVGTIGVSLLSNGLVLLCLLHSADIRRQAPALFTLNLTCGNLLCTVVNM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              | || ||:  | |:    |:: |||||||:|: ||:|||| |:|:|| ||| |  ::|
Db         61 PLTLAGVVAQRQPAGDRLCRLAAFLDTFLAANSMLSMAALSIDRWVAVVFPLSYRAKMRL 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              | |  ::   |  :| |   ||  ||||:   :|||:|       || ||| ||:  ||:
Db        121 RDAAFMVAYTWLHALTFPATALALSWLGFHQLYASCTLCSRRPDERLRFAVFTSAFHALS 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |:|   ||| | |:| :||| ||:|:| :||: | || |:|||||:||| :||||| |||
Db        181 FLLSFIVLCFTYLKVLKVARFHCKRIDVITMQTLVLLVDIHPSVRERCLEEQKRRRQRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              :||   | |||:||||||:||| ||     ::: ||:||||| |||| :||| |||||
Db        241 KKISTFIGTFLVCFAPYVITRLVELFSTAPIDSHWGVLSKCLAYSKAASDPFVYSLLRHQ 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVH 334
              :|:    :::|:  |     | |  |: : |  |
Db        301 YRRSCKELLNRIFNR----RSIH--SVGLTGDSH 328


RESULT 4
GP26_MOUSE
ID   GP26_MOUSE     STANDARD;      PRT;   337 AA.
AC   Q8BZA7; Q80T55; Q8BXZ5;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Probable G protein-coupled receptor GPR26.
GN   GPR26.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX   MEDLINE=22354683; PubMed=12466851;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
```

```
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [2]
RP   SEQUENCE OF 117-291 FROM N.A.
RX   MEDLINE=22584407; PubMed=12679517;
RA   Vassilatis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,
RA   Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,
RA   Bergmann J.E., Gaitanaris G.A.;
RT   "The G protein-coupled receptor repertoires of human and mouse.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908(2003).
CC   -!- FUNCTION: Orphan receptor.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Exclusively expressed in the brain.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AK036100; BAC29305.1; -.
DR   EMBL; AK042755; BAC31354.1; -.
DR   EMBL; AY255585; AAO85097.1; -.
DR   MGD; MGI:2441758; Gpr26.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR000169; SHprot_acsite.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane.
FT   DOMAIN        1     10       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     11     31       1 (POTENTIAL).
FT   DOMAIN       32     47       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     48     68       2 (POTENTIAL).
FT   DOMAIN       69     81       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     82    102       3 (POTENTIAL).
FT   DOMAIN      103    123       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    124    144       4 (POTENTIAL).
FT   DOMAIN      145    168       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    169    189       5 (POTENTIAL).
FT   DOMAIN      190    245       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    246    266       6 (POTENTIAL).
FT   DOMAIN      267    276       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    277    297       7 (POTENTIAL).
FT   DOMAIN      298    337       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     79    156       BY SIMILARITY.
FT   CONFLICT    156    156       C -> R (IN REF. 2).
SQ   SEQUENCE   337 AA;  37746 MW;  CD7CF15744E722DD CRC64;

  Query Match             44.2%;  Score 822;  DB 1;  Length 337;
  Best Local Similarity   49.1%;  Pred. No. 3.3e-54;
  Matches  164;  Conservative   57;  Mismatches  107;  Indels    6;  Gaps    2;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              |   :| ||||||  : |:|||| |||||  :||::| :|  :  :||: |:||   ::|

Db          1 MNSWDAGLAGLLVGTIGVSLLSNGLVLLCLLHSADIRRQAPALFTLNLTCGNLLCTVVNM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              | || ||:  | |:    |::  |||||||:|: ||:|||| |:|:|| ||| |  ::|

Db         61 PLTLAGVVAQRQPAGDRLCRLAAFLDTFLAANSMLSMAALSIDRWVAVVFPLSYRAKMRL 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              | |  ::   |  :| |   || ||||:   :|||:|      || ||| ||:  ||:

Db        121 RDAAFMVAYTWLHALTFPATALALSWLGFHQLYASCTLCSRRPDERLRFAVFTSAFHALS 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |:|   ||| | |:| :||| ||:|:| :||: | || |:|||||:||| :|||| |||

Db        181 FLLSFIVLCFTYLKVLKVARFHCKRIDVITMQTLVLLVDIHPSVRERCLEEQKRRRQRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              :||   | |||:||||||:||| ||     : : ||:||||| |||| :||| |||||

Db        241 KKISTFIGTFLVCFAPYVITRLVELFSTAPIGSHWGVLSKCLAYSKAASDPFVYSLLRHQ 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVH 334
              :|:    :::|:  |     | |  |: : |  |
Db        301 YRRSCKELLNRIFNR----RSLH--SVGLTGDSH 328


RESULT 5
5H6_MOUSE
ID   5H6_MOUSE      STANDARD;      PRT;   440 AA.
AC   Q9RIC8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   5-hydroxytryptamine 6 receptor (5-HT-6) (Serotonin receptor).
GN   HTR6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ; TISSUE=Brain;
RX   MEDLINE=21299097; PubMed=11406289;
RA   Kohen R., Fashingbauer L.A., Heidmann D.E.A., Guthrie C.R.,
RA   Hamblin M.W.;
RT   "Cloning of the mouse 5-HT6 serotonin receptor and mutagenesis
RT   studies of the third cytoplasmic loop.";
RL   Brain Res. Mol. Brain Res. 90:110-117(2001).
CC   -!- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR 5-
CC       HYDROXYTRYPTAMINE (SERONTONIN), A BIOGENIC HORMONE THAT FUNCTION AS
CC       A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF THIS
CC       RECEPTOR IS MEDIATED BY G PROTEINS THAT STIMULATES ADENYLATE
CC       CYCLASE. IT HAS A HIGH AFFINITY FOR TRICYCLIC PSYCHOTROPIC DRUGS
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF134158; AAD46490.1; -.
DR   MGD; MGI:1196627; Htr6.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family.
FT   DOMAIN        1     34       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     35     57       1 (POTENTIAL).
```

```
FT   DOMAIN       58     64       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     65     85       2 (POTENTIAL).
FT   DOMAIN       86    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    122       3 (POTENTIAL).
FT   DOMAIN      123    144       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    145    166       4 (POTENTIAL).
FT   DOMAIN      167    184       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    185    208       5 (POTENTIAL).
FT   DOMAIN      209    267       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    268    292       6 (POTENTIAL).
FT   DOMAIN      293    297       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    298    322       7 (POTENTIAL).
FT   DOMAIN      323    440       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     99    180       BY SIMILARITY.
FT   CARBOHYD      9      9       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   440 AA;  46998 MW;  4440CDEBE01FEF0C CRC64;

  Query Match             13.2%;  Score 245.5;  DB 1;  Length 440;
  Best Local Similarity   25.1%;  Pred. No. 3e-11;
  Matches  100;  Conservative   68;  Mismatches  174;  Indels   57;  Gaps   15;

Qy           2 GPGEA--------LLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHL 53
               |||          : | | |:::  |  :: |: | |  |   |   |   ||:|   |
Db          16 GPGPPPAPGGSGWVAAALCVVIVLTAAANSLLIALICTQPALRNT--SNFFLVSLFTSDL 73

Qy          54 LLAALDMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLR 113
               ::  : ||  :|  : ||    | | | :    |    | : |::  :| |::| :  |||
Db          74 MVGLVVMPPAMLNALYGRWVLARGLCLLWTAFDVMCCSASILNLCLISLDRYLLILSPLR 133

Qy         114 YAGRL-RPRYAGLLLGCAWGQSLA----FSGAALGCSWLG--YSSAFASCSL--RLPPEP 164
               |  |:  ||   |:|| ||  |||    |    ||   :||   | |   ||
Db         134 YKLRMTAPRALALILG-AW--SLAALASFLPLLLGWHELGKARTSAPGQCRLLASLP--- 187

Qy         165 ERPRFAAFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVT---------MKALA 215
                     :       | | ||   :| |  ::    ||:   :: ::|          :: |
Db         188 -------YVLVASGVTFFLPSGAICFTYCRILLAARKQAVQVASLTTGTATAGQALETLQ 240

Qy         216 LLADLHPSVR----QRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTV 271
               :    | :     :|  :  |:  :|:  :|| :: | : : |: :  :|: |
Db         241 VPRTPRPGMESADSRRLTTKHSRKALKASLTLGILLSMFFVTWLPFFVASIAQAVCDCIS 300

Qy         272 NAQWGILSKCLTYSKAVADPFTYSLLRRPFRQVLAGMV--------HRLLKRTPRPASTH 323
                 : :|:  | |  :  :|  | |  | |:: |   |          ||    :|   ::|
Db         301 PGLFDVLT-WLGYCNSTMNPIIYPLFMRDFKRALGRFVPCVHCPPEHRASPASPSMWTSH 359

Qy         324 DSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQQT 362
                 :    |:  | :   | | :: : |  :   | || |
Db         360 SGAR--PGLSLQQVLPLPLPPNSDSDSA-SGGTSGLQLT 395


RESULT 6
5H6_HUMAN
ID   5H6_HUMAN      STANDARD;      PRT;   440 AA.
AC   P50406; Q13640;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   5-hydroxytryptamine 6 receptor (5-HT-6) (Serotonin receptor).
GN   HTR6.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Striatum;
RX   MEDLINE=96102917; PubMed=8522988;
RA   Kohen R., Metcalf M.A., Khan N., Druck T., Huebner K., Lachowicz J.E.,
RA   Sibley D.R., Roth B., Hamblin M.W.;
RT   "Cloning, characterization, and chromosomal localization of a human
RT   5-HT6 serotonin receptor.";
RL   J. Neurochem. 66:47-56(1996).
RN   [2]
RP   SEQUENCE OF 215-280 FROM N.A.
RC   TISSUE=Striatum;
RX   MEDLINE=95385798; PubMed=7656980;
RA   Ullmer C., Schmuck K., Kalkman H.O., Lubbert H.;
RT   "Expression of serotonin receptor mRNAs in blood vessels.";
RL   FEBS Lett. 370:215-221(1995).
CC   -!- FUNCTION: This is one of the several different receptors for 5-
CC       hydroxytryptamine (serotonin), a biogenic hormone that functions
CC       as a neurotransmitter, a hormone, and a mitogen. The activity of
CC       this receptor is mediated by G proteins that stimulates adenylate
CC       cyclase. It has a high affinity for tricyclic psychotropic drugs.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in several human brain regions, most
CC       prominently in the caudate nucleus.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L41147; AAA92622.1; -.
DR   EMBL; Z49119; CAA88929.1; -.
DR   PIR; JC5520; JC5520.
DR   Genew; HGNC:5301; HTR6.
DR   MIM; 601109; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004969; F:histamine receptor activity; TAS.
DR   GO; GO:0007187; P:G-protein signaling, coupled to cyclic nucl. . .; TAS.
DR   GO; GO:0007268; P:synaptic transmission; TAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family.
FT   DOMAIN        1     34       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     35     57       1 (POTENTIAL).
FT   DOMAIN       58     64       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     65     85       2 (POTENTIAL).
FT   DOMAIN       86    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    101    122       3 (POTENTIAL).
FT   DOMAIN      123    144       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    145    166       4 (POTENTIAL).
FT   DOMAIN      167    184       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    185    208       5 (POTENTIAL).
FT   DOMAIN      209    265       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    266    290       6 (POTENTIAL).
FT   DOMAIN      291    295       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    296    320       7 (POTENTIAL).
FT   DOMAIN      321    440       CYTOPLASMIC (POTENTIAL).
FT   DISULFID     99    180       BY SIMILARITY.
FT   CONFLICT    247    247       V -> M (IN REF. 2).
SQ   SEQUENCE   440 AA;  46954 MW;  C888F47650C1D2EF CRC64;

  Query Match             13.2%;  Score 244.5;  DB 1;  Length 440;
  Best Local Similarity   26.1%;  Pred. No. 3.6e-11;
  Matches  102;  Conservative   64;  Mismatches  164;  Indels   61;  Gaps   17;

Qy           2 GP-----GEALLAGLLVMVLAVALLSNALVL-LCCAYSAELRTRASGVLLVNLSLGHLLL 55
                 ||     |   :|  | :|:|:   :|:|:: | |  |   |   |  ||:|    |::
Db          18 GPPSAPGGSGWVAAALCVVIALTAAANSLLIALICTQPALRNT--SNFFLVSLFTSDLMV 75

Qy          56 AALDMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYA 115
               : ||  :|  : ||   | | | :    |   |  : |::  :| |::| :  ||||
Db          76 GLVVMPPAMLNALYGRWVLARGLCLLWTAFDVMCCSASILNLCLISLDRYLLILSPLRYK 135

Qy         116 GRLRP-RYAGLLLGCAWGQSLA----FSGAALGCSWLGYSSAFASCSLRLPPEPERPRFA 170
                 |: | |   |:|| ||  |||    |    ||   ||::          || | : |
```

```
Db        136 LRMTPLRALALVLG-AW--SLAALASFLPLLLGWHELGHAR---------PPVPGQCRLL 183

Qy        171 A---FTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVT-------MKALALLADL 220

Db        184 ASLPFVLVASGLTFFLPSGAICFTYCRILLAARKQAVQVASLTTGMASQASETLQVPRTP 243

Qy        221 HPSVR----QRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWG 276

Db        244 RPGVESADSRRLATKHSRKALKASLTLGILLGMFFVTWLPFFVANIVQAVCDCISPGLFD 303

Qy        277 ILSKCLTYSKAVADPFTYSLLRRPFRQVLAGMVHRLL--KRTP--RPASTHDSSLDVA-- 330

Db        304 VLT-WLGYCNSTMNPIIYPLFMRDFKRALG----RFLPCPRCPRERQASLASPSLRTSHS 358

Qy        331 ----GMVHQLLKRTPRPASTHNGSVDTENDS 357

Db        359 GPRPGLSLQQVLPLPLPP-------DSDSDS 382


RESULT 7
SSR4_HUMAN
ID   SSR4_HUMAN     STANDARD;      PRT;   388 AA.
AC   P31391; Q9UIY1;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Somatostatin receptor type 4 (SS4R).
GN   SSTR4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93290656; PubMed=8512564;
RA   Xu Y., Song J., Bruno J.F., Berelowitz M.;
RT   "Molecular cloning and sequencing of a human somatostatin receptor,
RT   hSSTR4.";
RL   Biochem. Biophys. Res. Commun. 193:648-652(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93248256; PubMed=8483934;
RA   Rohrer L., Raulf F., Bruns C., Buettner R., Hofstaedter F.,
RA   Schuele R.;
RT   "Cloning and characterization of a fourth human somatostatin
RT   receptor.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:4196-4200(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93384611; PubMed=8373420;
RA   Yamada Y., Kagimoto S., Kubota A., Yasuda K., Masuda K., Someya Y.,
RA   Ihara Y., Li Q., Imura H., Seino S., Seino Y.;
RT   "Cloning, functional expression and pharmacological characterization
RT   of a fourth (hSSTR4) and a fifth (hSSTR5) human somatostatin receptor
RT   subtype.";
RL   Biochem. Biophys. Res. Commun. 195:844-852(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93302729; PubMed=8100352;
RA   Demchyshyn L.L., Srikant C.B., Sunahara R.K., Kent G., Seeman P.,
RA   van Tol H.H.M., Panetta R., Patel Y.C., Niznik H.B.;
RT   "Cloning and expression of a human somatostatin-14-selective receptor
RT   variant (somatostatin receptor 4) located on chromosome 20.";
RL   Mol. Pharmacol. 43:894-901(1993).
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21638749; PubMed=11780052;
RA   Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA   Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA   Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA   Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA   Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA   Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA   Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA   Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA   Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA   Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA   Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA   Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA   Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA   Lehvaeslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA   Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA   Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA   Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA   Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA   Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
RA   Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA   Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA   Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA   Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA   Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA   Rogers J.;
RT   "The DNA sequence and comparative analysis of human chromosome 20.";
RL   Nature 414:865-871(2001).
CC   -!- FUNCTION: Receptor for somatostatin-14. The activity of this
CC       receptor is mediated by G proteins which inhibits adenylyl
CC       cyclase. It is functionally coupled not only to inhibition of
CC       adenylate cyclase, but also to activation of both arachidonate
CC       release and mitogen-activated protein (MAP) kinase cascade.
CC       Mediates antiproliferative action of somatostatin in tumor cells.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Specifically expressed in fetal and adult
CC       brain, lung tissue, stomach, and in lesser quantities in the
CC       kidney, pituitary and adrenals.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D16826; BAA04106.1; -.
DR   EMBL; L14856; AAA36623.1; -.
DR   EMBL; L07833; AAA60565.1; -.
DR   EMBL; L07061; -; NOT_ANNOTATED_CDS.
DR   EMBL; AL049651; CAB51953.1; -.
DR   PIR; JN0605; JN0605.
DR   HSSP; P02699; 1BOJ.
DR   Genew; HGNC:11333; SSTR4.
DR   MIM; 182454; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004994; F:somatostatin receptor activity; TAS.
DR   GO; GO:0007187; P:G-protein signaling, coupled to cyclic nucl. . .; TAS.
DR   GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Lipoprotein; Palmitate; Phosphorylation;
KW   Polymorphism.
FT   DOMAIN       1     46       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    47     73       1 (POTENTIAL).
FT   DOMAIN      74     83       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    84    104       2 (POTENTIAL).
FT   DOMAIN     105    120       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   121    142       3 (POTENTIAL).
FT   DOMAIN     143    161       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   162    185       4 (POTENTIAL).
FT   DOMAIN     186    208       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   209    232       5 (POTENTIAL).
FT   DOMAIN     233    260       CYTOPLASMIC (POTENTIAL).
```

```
FT   TRANSMEM     261    280       6 (POTENTIAL).
FT   DOMAIN       281    291       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     292    314       7 (POTENTIAL).
FT   DOMAIN       315    388       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      24     24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     119    198       BY SIMILARITY.
FT   LIPID        327    327       S-palmitoyl cysteine (Potential).
FT   VARIANT       83     83       N -> T (in dbSNP:1065191).
FT                                 /FTId=VAR_011703.
FT   CONFLICT     284    284       V -> F (IN REF. 4 AND 5).
FT   CONFLICT     321    321       S -> F (IN REF. 5).
FT   CONFLICT     365    365       P -> K (IN REF. 2).
SQ   SEQUENCE   388 AA;  41894 MW;  8BCCD69B5F3BC2F5 CRC64;

  Query Match             13.1%;  Score 243.5;  DB 1;  Length 388;
  Best Local Similarity   25.5%;  Pred. No. 3.8e-11;
  Matches   84;  Conservative   55;  Mismatches  124;  Indels   67;  Gaps   10;

Qy          2 GPGEALLAGLL------VMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLL 55
              |||:|  ||::       :|  | |: ||||:       |:::| |: : |:||::   |
Db         38 GPGDARAAGMVAIQCIYALVCLVGLVGNALVIFVILRYAKMKT-ATNIYLLNLAVADELF 96

Qy         56 AALDMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYA 115
                | :||          |     |: :  :|       :   :  || |:::||  ||| |
Db         97 -MLSVPFVASSAALRHWPFGSVLCRAVLSVDGLNMFTSVFCLTVLSVDRYVAVVHPLRAA 155

Qy        116 GRLRPRYAGLLLGCAWGQSL-------------AFSGAALGCS--WLGYSSAFASCSLR 159
                ||  | |:     |  ||               |  | |: |:  |
Db        156 TYRRPSVAKLINLGVWLASLLVTLPIAIFADTRPARGGQAVACNLQW------------ 202

Qy        160 LPPEPERPRFAAFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLAD 219
                |      | |       :||:||:  : |           |  :    |:|:|| |
Db        203 ----PHPAWSAVFVVYTFLLGFLLPVLAIGL-----------CYLLIVGKMRAVALRAG 246

Qy        220 LHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILS 279
                    |::|  : || : : :  |::|: |: : :|  ||   :::|   :|
Db        247 WQ---------QRRRSEKKITRLVLMVVVVFVLCWMPFYVVQLLNLV-VTSLDATVNHVS 296

Qy        280 KCLTYSKAVADPFTYSLL----RRPFRQVL 305
                |:|: : |:|  |  |    || |::||
Db        297 LILSYANSCANPILYGFLSDNFRRSFQRVL 326


RESULT 8
5H6_RAT
ID   5H6_RAT        STANDARD;      PRT;   436 AA.
AC   P31388;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   5-hydroxytryptamine 6 receptor (5-HT-6) (Serotonin receptor)
DE   (ST-B17).
GN   HTR6.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Striatum;
RX   MEDLINE=93196608; PubMed=7680751;
RA   Monsma F.J. Jr., Shen Y., Ward R.P., Hamblin M.W., Sibley D.R.;
RT   "Cloning and expression of a novel serotonin receptor with high
RT   affinity for tricyclic psychotropic drugs.";
RL   Mol. Pharmacol. 43:320-327(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93277562; PubMed=8389146;
RA   Ruat M., Traiffort E., Arrang J.-M., Tardivel-Lacombe J., Diaz J.,
RA   Leurs R., Schwartz J.-C.;
RT   "A novel rat serotonin (5-HT6) receptor: molecular cloning,
RT   localization and stimulation of cAMP accumulation.";
RL   Biochem. Biophys. Res. Commun. 193:268-276(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar;
RA   Martial R.;
RL   Submitted (XXX-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: This is one of the several different receptors for 5-
CC       hydroxytryptamine (serotonin), a biogenic hormone that functions
CC       as a neurotransmitter, a hormone, and a mitogen. The activity of
CC       this receptor is mediated by G proteins that stimulates adenylate
CC       cyclase. It has a high affinity for tricyclic psychotropic drugs.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Localized exclusively in the central nervous
CC       system, predominantly in the corpus striatum but also in various
CC       limbic and cortical regions.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L03202; AAA40618.1; -.
DR   EMBL; S62043; AAB26908.1; -.
DR   EMBL; L19656; AAA40611.1; -.
DR   PIR; I57942; I57942.
DR   PIR; JN0591; JN0591.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family.
FT   DOMAIN         1     34       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM      35     57       1 (POTENTIAL).
FT   DOMAIN        58     64       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      65     85       2 (POTENTIAL).
FT   DOMAIN        86    100       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     101    122       3 (POTENTIAL).
FT   DOMAIN       123    144       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     145    166       4 (POTENTIAL).
FT   DOMAIN       167    184       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     185    208       5 (POTENTIAL).
FT   DOMAIN       209    265       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     266    290       6 (POTENTIAL).
FT   DOMAIN       291    295       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     296    320       7 (POTENTIAL).
FT   DOMAIN       321    436       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD       9      9       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID      99    180       BY SIMILARITY.
FT   CONFLICT      57     57       L -> V (IN REF. 1).
FT   CONFLICT     336    436       PCVHCPPEHRPALPPPPCGPLTAVPDQASACSRCCLCLCRQ
FT                                 TQIQTPLQGAPRACSSQPSFCCLERPPGTPRHPPGPPLWST
FT                                 SLSQTLWSLRYGRIHSVPP -> HASTVPRSTGQPCLPLHV
FT                                 DLSQRCQTRPQLQQVLALPLPPNSDSDSASGGTSGLQLTAQ
FT                                 LLLPGEATRDPPPPPTRATTVVNFFVTDSVEPEIRPHPLSSP
FT                                 VN (IN REF. 1).
SQ   SEQUENCE   436 AA;  46922 MW;  6BC00F6A3CBA5FB4 CRC64;

  Query Match             12.9%;  Score 239;  DB 1;  Length 436;
  Best Local Similarity   25.7%;  Pred. No. 9.2e-11;
  Matches   85;  Conservative   58;  Mismatches  146;  Indels   42;  Gaps   11;

Qy          2 GPGEA--------LLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHL 53
              |||          : | | |:::  |  :: |::|  |    |   |   ||:|    |
Db         16 GPGPPPAPGGSGWVAAALCVVIVLTAAANSLLIVLICTQPALRNT--SNFFLVSLFTSDL 73

Qy         54 LLAALDMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLR 113
              ::  : ||  :|  : ||    | | | :    |    | : |::  :| |::| :  |||
Db         74 MVGLVVMPPAMLNALYGRWVLARGLCLLWTAFDVMCCSASILNLCLISLDRYLLILSPLR 133
```

```
Qy        114 YAGRL-RPRYAGLLLGCAWGQSLA----FSGAALGCSWLGYSSAFASCSLRLPPEPERPR 168
Db        134 YKLRMTAPRALALILG-AW--SLAALASFLPLLLGWHELGKART---------PAPGQCR 181

Qy        169 FAA---FTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVT-------MKALALLA 218
Db        182 LLASLPFVLVASGVTFFLPSGAICFTYCRILLAARKQAVQVASLTTGTAGQALETLQVPR 241

Qy        219 DLHPSVR----QRCLIQQKRRRHRATRKIGIAIATFLICFAPYVVMTRLAELVPFVTVNAQ 274
Db        242 TPRPGMESADSRRLATKHSRKALKASLTLGILLGMFFVTWLPFFVANIAQAVCDCISPGL 301

Qy        275 WGILSKCLTYSKAVADPFTYSLLRRPFRQVL 305
Db        302 FDVLT-WLGYCNSTMNPIIYPLFMRDFKRAL 331
```

```
RESULT 9
HH2R_CAVPO
ID   HH2R_CAVPO     STANDARD;      PRT;   359 AA.
AC   P47747;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Histamine H2 receptor (H2R) (Gastric receptor I).
GN   HRH2.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Hartley; TISSUE=Liver;
RA   Traiffort E.;
RL   Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: The H2 subclass of histamine receptors mediates gastric
CC       acid secretion. The activity of this receptor is mediated by G
CC       proteins which activate adenylyl cyclase.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U25440; AAA65713.1; -.
DR   PIR; JC4120; JC4120.
DR   HSSP; P29274; 1MMH.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein;
KW   Palmitate.
FT   DOMAIN        1     22       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     23     44       1 (POTENTIAL).
FT   DOMAIN       45     57       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     58     81       2 (POTENTIAL).
FT   DOMAIN       82     92       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     93    114       3 (POTENTIAL).
FT   DOMAIN      115    134       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    135    159       4 (POTENTIAL).
FT   DOMAIN      160    180       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    181    204       5 (POTENTIAL).
FT   DOMAIN      205    234       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    235    258       6 (POTENTIAL).
FT   DOMAIN      259    267       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    268    289       7 (POTENTIAL).
FT   DOMAIN      290    359       CYTOPLASMIC (POTENTIAL).
FT   SITE         98     98       ESSENTIAL FOR HISTAMINE BINDING (BY
FT                                SIMILARITY).
FT   SITE        186    186       ESSENTIAL FOR TIOTIDINE BINDING AND
FT                                IMPLICATED IN HISTAMINE BINDING (BY
FT                                SIMILARITY).
FT   SITE        190    190       IMPLICATED IN HISTAMINE BINDING (BY
FT                                SIMILARITY).
FT   LIPID       305    305       S-palmitoyl cysteine (By similarity).
FT   CARBOHYD      4      4       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     91    174       BY SIMILARITY.
SQ   SEQUENCE   359 AA;  40556 MW;  58DB81BD8FC3C0E9 CRC64;

  Query Match             12.5%;  Score 232.5;  DB 1;  Length 359;
  Best Local Similarity   22.0%;  Pred. No. 2.3e-10;
  Matches   73;  Conservative   67;  Mismatches  139;  Indels   53;  Gaps    8;

Qy         11 LLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGVMRG 70
Db         24 ILIILILVTVAGNVVVCLAVGLNRRLRS-LTNCFIVSLAVTDLLLGLLVLPFSAIYQLSC 82

Qy         71 RTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLLLGCA 130
Db         83 KWSFSKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPARVAISLVFI 142

Qy        131 WGQSLAFSGAALGCSWLGY------SSAFASCSLRLPPEPERPRFAAFTATLHAVGFVLP 184
Db        143 WVISITLSFLSIHLGWNSRNETSKDNDTIVKCKVQVN--------EVYGLVDGLVTFYLP 194

Qy        185 LAVLCLTSLQVHRVARRHCQRMDTV-TMKALALLADLHPSVRQRCLIQQKRRRHRATRKI 243
Db        195 LLIMCITYFRIFKIAREQARRINHIGSWKAATI-----------------REHKATVTL 236

Qy        244 GIAIATFLICFAPYVMTRLAELVPFV--------TVNAQWGILSKCLTYSKAVADPFTYS 295
Db        237 AAVMGAFIICWFPY-------FTVFVYRGLKGDDAVNEVFEDVVLWLGYANSALNPILYA 289

Qy        296 LLRRPFRQVLAGMVHRLLKRTPRPASTHDSSL 327
Db        290 ALNRDFRTA----YHQLFCCRLASHNSHETSL 317
```

```
RESULT 10
5H4_MOUSE
ID   5H4_MOUSE      STANDARD;      PRT;   388 AA.
AC   P97288; O89003; O89004; Q9R2A4;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   5-hydroxytryptamine 4 receptor (5-HT-4) (Serotonin receptor) (5-HT4).
GN   HTR4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss; TISSUE=Brain;
RX   MEDLINE=97102706; PubMed=8946946;
RA   Claeysen S., Sebben M., Journot L., Bockaert J., Dumuis A.;
RT   "Cloning, expression and pharmacology of the mouse 5-HT(4L)
RT   receptor.";
RL   FEBS Lett. 398:19-25(1996).
RN   [2]
RP   REVISIONS TO C-TERMINUS.
RA   Dumuis A.;
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS 5-HT4(A) AND 5-HT4(E)).
RC   STRAIN=Swiss; TISSUE=Brain;
RX   MEDLINE=99127199; PubMed=9928238;
RA   Claeysen S., Faye P., Sebben M., Taviaux S., Bockaert J., Dumuis A.;
```

```
RT   "5-HT4 receptors: cloning and expression of new splice variants.";
RL   Ann. N.Y. Acad. Sci. 861:49-56(1998).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 5-HT4(F)).
RC   TISSUE=Brain;
RX   MEDLINE=99238795; PubMed=10220570;
RA   Claeysen S., Sebben M., Becamel C., Bockaert J., Dumuis A.;
RT   "Novel brain-specific 5-HT4 receptor splice variants show marked
RT   constitutive activity: role of the C-terminal intracellular domain.";
RL   Mol. Pharmacol. 55:910-920(1999).
CC   -!- FUNCTION: This is one of the several different receptors for 5-
CC       hydroxytryptamine (serotonin), a biogenic hormone that functions
CC       as a neurotransmitter, a hormone, and a mitogen. The activity of
CC       this receptor is mediated by G proteins that stimulates adenylate
CC       cyclase (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=1;
CC         IsoId=P97288-1; Sequence=Displayed;
CC       Name=5-HT4(A);
CC         IsoId=P97288-2; Sequence=VSP_001851;
CC       Name=5-HT4(E);
CC         IsoId=P97288-3; Sequence=VSP_001852;
CC       Name=5-HT4(F);
CC         IsoId=P97288-4; Sequence=VSP_001853;
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y09585; CAA70773.1; -.
DR   EMBL; Y09587; CAA70775.1; -.
DR   EMBL; Y09588; CAA70776.1; -.
DR   EMBL; AJ011369; CAA09598.1; -.
DR   HSSP; P29274; 1MMH.
DR   MGD; MGI:109246; Htr4.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Lipoprotein; Palmitate; Alternative splicing.
FT   DOMAIN        1     19       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     20     40       1 (POTENTIAL).
FT   DOMAIN       41     58       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     59     79       2 (POTENTIAL).
FT   DOMAIN       80     93       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     94    116       3 (POTENTIAL).
FT   DOMAIN      117    137       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    138    158       4 (POTENTIAL).
FT   DOMAIN      159    192       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    193    213       5 (POTENTIAL).
FT   DOMAIN      214    260       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    261    281       6 (POTENTIAL).
FT   DOMAIN      282    294       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    295    315       7 (POTENTIAL).
FT   DOMAIN      316    388       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      7      7       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID     93    184       BY SIMILARITY.
FT   LIPID       329    329       S-palmitoyl cysteine (By similarity).
FT   VARSPLIC    360    388       DAVECGGQWESRCHLTATSPLVAAQPSDT -> YTVLHSGH
FT                                HQELEKLPIHNDPESLESCF (in isoform
FT                                5-HT4(A)).
FT                                /FTId=VSP_001851.
FT   VARSPLIC    359    388       RDAVECGGQWESRCHLTATSPLVAAQPSDT -> SFPLLFR
FT                                NRPVPV (in isoform 5-HT4(E)).
FT                                /FTId=VSP_001852.
FT   VARSPLIC    360    388       DAVECGGQWESRCHLTATSPLVAAQPSDT -> PVPV (in
FT                                isoform 5-HT4(F)).
FT                                /FTId=VSP_001853.
SQ   SEQUENCE   388 AA;  43931 MW;  F84163F39BA115B0 CRC64;

  Query Match             12.5%;  Score 232;  DB 1;  Length 388;
  Best Local Similarity   22.1%;  Pred. No. 2.7e-10;
  Matches   83;  Conservative   72;  Mismatches  166;  Indels   54;  Gaps   10;

Qy          5 EALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTL 64
              : :|   | :|: :|:| | ||::      :||   :   :|:|:   ||:: | |||
Db         19 KVVLLTFLAVVILMAILGNLLVMVAVCRDRQLRKIKTNYFIVSLAFADLLVSVLVMPFGA 78

Qy         65 LGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGF-PLRYAGRLRPRYA 123
              : :::         | |   ||  | : :   :  :| |:: |:   || |  :: |
Db         79 IELVQDIWAYGEMFCLVRTSLDVLLTTASIFHLCCISLDRYYAICCQPLVYRNKMTPLRI 138

Qy        124 GLLLGCAWGQSLAFSGAALGCSW----------LGYSSAFASCSLRLPPEPERPRFA 170
              |:|| |   :  |   :  |             ::|    |   :    :|
Db        139 ALMLGGCWVLPMFISFLPIMQGWNNIGIVDVIEKRKFSHNSNSTWCVFMV----NKP--- 191

Qy        171 AFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTV-----TMKALALLADLHPSVR 225
              :  |   | | :|  :: |   :::  |: | |::  :      | ::    || | : |
Db        192 -YAITCSVVAFYIPFLLMVLAYYRIYVTAKEHAQQIQMLQRAGATSESRPQPADQHSTHR 250

Qy        226 QRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFV--TVNAQWGILSKCLT 283
                  |  :| : : : :  |  |:||: :| : :  ||:  ||  |          |
Db        251 M-------RTETKAAKTLCVIMGCFCFCWAPFFVTNIVD--PFIDYTVPEQVWTAFLWLG 301

Qy        284 YSKAVADPFTYSLLRRPFRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRT-PR 342
              |  :  :|| |: | : ||:    ::    :|  ||              :| :| | |
Db        302 YINSGLNPFLYAFLNKSFRRAFLIILCCDDERYKRP---------------PILGQTVPC 346

Qy        343 PASTHNGSVDTENDS 357
              :| |||      |:
Db        347 STTTINGSTHVLRDA 361


RESULT 11
SSR5_HUMAN
ID   SSR5_HUMAN     STANDARD;      PRT;   364 AA.
AC   P35346; P34988; Q9UJI5;
DT   01-JUN-1994 (Rel. 29, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Somatostatin receptor type 5 (SS5R).
GN   SSTR5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94195267; PubMed=7908405;
RA   Panetta R., Greenwood M.T., Warszynska A., Demchyshyn L.L., Day R.,
RA   Niznik H.B., Srikant C.B., Patel Y.C.;
RT   "Molecular cloning, functional characterization, and chromosomal
RT   localization of a human somatostatin receptor (somatostatin receptor
RT   type 5) with preferential affinity for somatostatin-28.";
RL   Mol. Pharmacol. 45:417-427(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93384611; PubMed=8373420;
RA   Yamada Y., Kagimoto S., Kubota A., Yasuda K., Masuda K., Someya Y.,
RA   Ihara Y., Li Q., Imura H., Seino S., Seino Y.;
RT   "Cloning, functional expression and pharmacological characterization
RT   of a fourth (hSSTR4) and a fifth (hSSTR5) human somatostatin receptor
RT   subtype.";
RL   Biochem. Biophys. Res. Commun. 195:844-852(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94359492; PubMed=8078491;
```

```
RA   O'Carroll A.-M., Raynor K., Lolait S.J., Reisine T.;
RT   "Characterization of cloned human somatostatin receptor SSTR5.";
RL   Mol. Pharmacol. 46:291-298(1994).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21096910; PubMed=11157797;
RA   Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA   Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA   Higgs D.R.;
RT   "Sequence, structure and pathology of the fully annotated terminal 2
RT   Mb of the short arm of human chromosome 16.";
RL   Hum. Mol. Genet. 10:339-352(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Bagguley C.;
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for somatostatin 28 and to a lesser extent for
CC       somatostatin-14. The activity of this receptor is mediated by G
CC       proteins which inhibit adenylyl cyclase.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Adult pituitary gland, heart, small intestine,
CC       adrenal gland, cerebellum and fetal hypothalamus. No expression
CC       in fetal or adult kidney, liver, pancreas, uterus, spleen, lung,
CC       thyroid or ovary.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14865; AAA20828.1; -.
DR   EMBL; D16827; BAA04107.1; -.
DR   EMBL; AE006466; AAK61266.1; -.
DR   EMBL; AL031713; CAB56181.1; -.
DR   PIR; I57955; I57955.
DR   PIR; JN0763; JN0763.
DR   HSSP; P02699; 1F88.
DR   Genew; HGNC:11334; SSTR5.
DR   MIM; 182455; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004994; F:somatostatin receptor activity; TAS.
DR   GO; GO:0007187; P:G-protein signaling, coupled to cyclic nucl. . .; TAS.
DR   GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Lipoprotein; Palmitate; Phosphorylation.
FT   DOMAIN        1     38       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     39     66       1 (POTENTIAL).
FT   DOMAIN       67     76       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     77    101       2 (POTENTIAL).
FT   DOMAIN      102    113       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    114    135       3 (POTENTIAL).
FT   DOMAIN      136    157       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    158    178       4 (POTENTIAL).
FT   DOMAIN      179    197       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    198    222       5 (POTENTIAL).
FT   DOMAIN      223    247       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    248    273       6 (POTENTIAL).
FT   DOMAIN      274    283       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    284    308       7 (POTENTIAL).
FT   DOMAIN      309    364       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     13     13       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    187    187       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    112    186       BY SIMILARITY.
FT   LIPID       320    320       S-palmitoyl cysteine (Potential).
FT   MOD_RES     325    325       PHOSPHORYLATION (BY PKA) (POTENTIAL).
FT   CONFLICT    335    335       P -> L (IN REF. 4 AND 5).
FT   CONFLICT    348    352       PPAHR -> RPRT (IN REF. 1).
SQ   SEQUENCE   364 AA;  39202 MW;  905744715F31121C CRC64;

  Query Match             12.5%;  Score 231.5;  DB 1;  Length 364;
  Best Local Similarity   25.0%;  Pred. No. 2.8e-10;
  Matches   92;  Conservative   51;  Mismatches  166;  Indels   59;  Gaps   11;

Qy          2 GPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMP 61
              |    |:  | ::| |   |  | ||:      |:::|  : : ::||::  :|   | :|
Db         37 GARAVLVPVLYLLVCAAGLGGNTLVIYVVLRFAKMKT-VTNIYILNLAVADVLY-MLGLP 94

Qy         62 FTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPR 121
              |          |  |  |:::  ||       :   :  :| |::|||  ||  |   |||
Db         95 FLATQNAASFWPFGPVLCRLVMTLDGVNQFTSVFCLTVMSVDRYLAVVHPLSSARWRRPR 154

Qy        122 YAGLLLGCAWGQSLAFS-----------GAALGCSWLGYSSAFASCSLRLPPEPERPRFA 170
              | |    ||  ||  |           |   ||                |||     |
Db        155 VAKLASAAAWVLSLCMSLPLLVFADVQEGGTCNASW--------------PEPVGLWGA 199

Qy        171 AFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLI 230
              |      :||  || |:||            |  :  | ::|          ||  |:
Db        200 VFIIYTAVLGFFAPLLVICL-----------CYLLIVVKVRA--------AGVRVGCV- 238

Qy        231 QQKRRRHRATRKIGIAIATFLICFAPYV---MTRLAELVPFVTVNAQWGILSKCLTYSKA 287
               ::|   : || : : :  |  |: |:    :  ||  :|    :|        |:|: :
Db        239 -RRRSERKVTRMVLVVVLVFAGCWLPFFTVNIVNLAVALPQEPASAGLYFFVVILSYANS 297

Qy        288 VADPFTYSLLRRPFRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPA--S 345
              |:|  |  |   |||    :   |  |    |   |::      : |  : || ||  :
Db        298 CANPVLYGFLSDNFRQSFQKV---LCLRKGSGAKDADATEPRPDRIRQQQEATP-PAHRA 353

Qy        346 THNGSVDT 353
                || : |
Db        354 AANGLMQT 361


RESULT 12
B3AR_HUMAN
ID   B3AR_HUMAN     STANDARD;      PRT;   408 AA.
AC   P13945;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Beta-3 adrenergic receptor.
GN   ADRB3 OR ADRB3R OR B3AR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89368947; PubMed=2570461;
RA   Emorine L.J., Marullo S., Briend-Sutren M.-M., Patey G., Tate K.,
RA   Delavier-Klutchko C., Strosberg A.D.;
RT   "Molecular characterization of the human beta 3-adrenergic receptor.";
RL   Science 245:1118-1121(1989).
RN   [2]
RP   REVISIONS, SEQUENCE FROM N.A.
RX   MEDLINE=93279311; PubMed=8389293;
RA   van Spronsen A., Nahmias C., Krief S., Briend-Sutren M.-M.,
RA   Strosberg A.D., Emorine L.J.;
RT   "The promoter and intron/exon structure of the human and mouse beta
RT   3-adrenergic-receptor genes.";
RL   Eur. J. Biochem. 213:1117-1124(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93285320; PubMed=8389717;
RA   Lelias J.M., Kaghad M., Rodriguez M., Chalon P., Bonnin J.,
RA   Dupre I., Delpech B., Bensaid M., Lefur G., Ferrara P., Caput D.;
RT   "Molecular cloning of a human beta 3-adrenergic receptor cDNA.";
```

```
RL   FEBS Lett. 324:127-130(1993).
RN   [4]
RP   SEQUENCE OF 392-408 FROM N.A.
RX   MEDLINE=93125503; PubMed=1336117;
RA   Granneman J.G., Lahners K.N., Rao D.D.;
RT   "Rodent and human beta 3-adrenergic receptor genes contain an intron
RT   within the protein-coding block.";
RL   Mol. Pharmacol. 42:964-970(1992).
RN   [5]
RP   VARIANT ARG-64.
RX   MEDLINE=95334045; PubMed=7609752;
RA   Clement K., Vaisse C., Manning B.S.J., Basdevant A., Guy-Grand B.,
RA   Ruiz J., Silver K.D., Shuldiner A.R., Froguel P., Strosberg A.D.;
RT   "Genetic variation in the beta 3-adrenergic receptor and an increased
RT   capacity to gain weight in patients with morbid obesity.";
RL   New Engl. J. Med. 333:352-354(1995).
RN   [6]
RP   VARIANT ARG-64.
RX   MEDLINE=96361936; PubMed=8721782;
RA   Fujisawa T., Ikegami H., Yamato E., Takekawa K., Nakagawa Y.,
RA   Hamada Y., Oga T., Ueda H., Shintani M., Fukuda M., Ogihara T.;
RT   "Association of Trp64Arg mutation of the beta3-adrenergic-receptor
RT   with NIDDM and body weight gain.";
RL   Diabetologia 39:349-352(1996).
RN   [7]
RP   VARIANT ARG-64.
RX   MEDLINE=96217306; PubMed=8641219;
RA   Candelore M.R., Deng L., Tota L.M., Kelly L.J., Cascieri M.A.,
RA   Strader C.D.;
RT   "Pharmacological characterization of a recently described human beta
RT   3-adrenergic receptor mutant.";
RL   Endocrinology 137:2638-2641(1996).
RN   [8]
RP   VARIANTS ARG-64 AND MET-265.
RX   MEDLINE=99318094; PubMed=10391210;
RA   Halushka M.K., Fan J.-B., Bentley K., Hsie L., Shen N., Weder A.,
RA   Cooper R., Lipshutz R., Chakravarti A.;
RT   "Patterns of single-nucleotide polymorphisms in candidate genes for
RT   blood-pressure homeostasis.";
RL   Nat. Genet. 22:239-247(1999).
CC   -!- FUNCTION: Beta-adrenergic receptors mediate the catecholamine-
CC       induced activation of adenylate cyclase through the action of G
CC       proteins. Beta-3 is involved in the regulation of lipolysis and
CC       thermogenesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed mainly in adipose tissues.
CC   -!- POLYMORPHISM: The variant Arg-64 seems to be associated with
CC       weight gain (obesity) and to is also associated with
CC       susceptibility to non-insulin-dependent diabetes mellitus (NIDDM).
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M29932; AAA35550.1; ALT_TERM.
DR   EMBL; X72861; CAA51383.1; -.
DR   EMBL; X70811; CAA50141.1; -.
DR   EMBL; X70812; CAA50142.1; ALT_SEQ.
DR   EMBL; X70812; CAA50143.1; ALT_SEQ.
DR   EMBL; S53291; AAB24837.1; -.
DR   PIR; I57941; I57941.
DR   PIR; S33751; QRHUBE.
DR   HSSP; P29274; 1MMH.
DR   Genew; HGNC:288; ADRB3.
DR   MIM; 109691; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0015052; F:beta3-adrenergic receptor activity; TAS.
DR   GO; GO:0005975; P:carbohydrate metabolism; TAS.
DR   GO; GO:0006091; P:energy pathways; TAS.
DR   GO; GO:0006112; P:energy reserve metabolism; TAS.
DR   GO; GO:0007188; P:G-protein signaling, coupled to cAMP nucleo. . .; TAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Phosphorylation; Lipoprotein; Palmitate;
KW   Polymorphism; Diabetes mellitus; Obesity.
FT   DOMAIN        1     36       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     37     63       1 (POTENTIAL).
FT   DOMAIN       64     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     73     91       2 (POTENTIAL).
FT   DOMAIN       92    111       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    112    133       3 (POTENTIAL).
FT   DOMAIN      134    155       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    156    178       4 (POTENTIAL).
FT   DOMAIN      179    203       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    204    225       5 (POTENTIAL).
FT   DOMAIN      226    292       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    293    314       6 (POTENTIAL).
FT   DOMAIN      315    326       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    327    347       7 (POTENTIAL).
FT   DOMAIN      348    408       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    110    189       BY SIMILARITY.
FT   LIPID       361    361       S-palmitoyl cysteine (By similarity).
FT   VARIANT      64     64       W -> R (in dbSNP:4994).
FT                                /FTId=VAR_003456.
FT   VARIANT     265    265       T -> M (in dbSNP:4995).
FT                                /FTId=VAR_014166.
SQ   SEQUENCE   408 AA;  43519 MW;  E98BD6C130DD977B CRC64;

  Query Match             12.4%;  Score 230.5;  DB 1;  Length 408;
  Best Local Similarity   24.9%;  Pred. No. 3.7e-10;
  Matches   93;  Conservative   61;  Mismatches  154;  Indels   65;  Gaps   14;

Qy          3 PGEALLAGLLVMVLAVALLS----NALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAAL 58
              | || ||| |   ||:|:|:    | ||::  |::  |:|  : | : :|:   |::  |
Db         34 PWEAALAGAL---LALAVLATVGGNLLVIVAIAWTPRLQT-MTNVFVTSLAAADLVMGLL 89

Qy         59 DMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRL 118
                :|      : |  |      |::  :|    : : :: ||: |::|||  |||   :
Db         90 VVPPAATLALTGHWPLGATGCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALV 149

Qy        119 RPRYAGLLLGCAWGQSLAFSGAALGCSW--LGYSSAFASCSLRLPPEPERPRFAAFTATL 176
              | |   :  |  | | | | | :   |  :| :    |          || || : :
Db        150 TKRCARTAVVLVWVVSAAVSFAPIMSQWWRVGADAEAQRCH-------SNPRCCAFASNM 202

Qy        177 ------HAVGFVLPLAVLCLTSLQVHRVARRHC------------QRMDTVTMKALA--- 215
                  :| | ||| |:     :|  || |              :       ::||
Db        203 PYVLLSSSVSFYLPLLVMLFVYARVFVVATRQLRLLRGELGRFPPEESPPAPSRSLAPAP 262

Qy        216 ----LLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRL------AEL 265
                  :  |:  :|       | |||   :|: : || :|: |: :  :        |
Db        263 VGTCAPPEGVPACGRRPARLLPLREHRALCTLGLIMGTFTLCWLPFFLANVLRALGGPSL 322

Qy        266 VP---FVTVNAQWGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTPRP 319
              ||   |: :|  |       | |: :  :|  |       |  ||::|     ||    | |
Db        323 VPGPAFLALN--W------LGYANSAFNPLIYCRSPDFRSAFRRLLCRCGRRL---PPEP 371

Qy        320 ASTHDSSLDVAGM 332
                :    :|  :|:
Db        372 CAAARPALFPSGV 384


RESULT 13
B3AR_MACMU
ID   B3AR_MACMU     STANDARD;      PRT;   418 AA.
AC   Q28524;
```

```
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Beta-3 adrenergic receptor.
GN   ADRB3 OR B3AR.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97279016; PubMed=9133593;
RA   Walston J., Lowe A., Silver K., Yang Y., Bodkin N.L., Hansen B.C.,
RA   Shuldiner A.R.;
RT   "The beta3-adrenergic receptor in the obesity and diabetes prone
RT   rhesus monkey is very similar to human and contains arginine at codon
RT   64.";
RL   Gene 188:207-213(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Thompson G.M., Kelly L.J., Candelore M.R.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Beta-adrenergic receptors mediate the catecholamine-
CC       induced activation of adenylate cyclase through the action of G
CC       proteins. Beta-3 is involved in the regulation of lipolysis and
CC       thermogenesis.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U63592; AAB53939.1; -.
DR   EMBL; U63591; AAB53939.1; JOINED.
DR   EMBL; AF200596; AAF08306.1; -.
DR   PIR; G02953; G02953.
DR   HSSP; P29274; 1MMH.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Phosphorylation; Lipoprotein; Palmitate.
FT   DOMAIN        1     36       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     37     63       1 (POTENTIAL).
FT   DOMAIN       64     72       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     73     91       2 (POTENTIAL).
FT   DOMAIN       92    111       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    112    133       3 (POTENTIAL).
FT   DOMAIN      134    155       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    156    178       4 (POTENTIAL).
FT   DOMAIN      179    203       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    204    225       5 (POTENTIAL).
FT   DOMAIN      226    292       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    293    314       6 (POTENTIAL).
FT   DOMAIN      315    326       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    327    347       7 (POTENTIAL).
FT   DOMAIN      348    418       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD      8      8       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     26     26       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    110    189       BY SIMILARITY.
FT   LIPID       361    361       S-palmitoyl cysteine (By similarity).
SQ   SEQUENCE   418 AA;  44657 MW;  F076D520ADC3502D CRC64;

  Query Match             12.4%;  Score 229.5;  DB 1;  Length 418;
  Best Local Similarity   25.9%;  Pred. No. 4.5e-10;

  Matches   95;  Conservative   56;  Mismatches  149;  Indels   67;  Gaps   15;

Qy          3 PGEALLAGLLVMVLAVALLS----NALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAAL 58
              |  | ||| |   ||:|:|:    | ||::    :  |:|  : | : :|:   |:: |
Db         34 PWAAALAGAL---LALAVLATVGGNLLVIVAITRTPRLQT-MTNVFVTSLAAADLVMGLL 89

Qy         59 DMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRL 118
              :|     |: |  |      |::   :|    : :  :: ||: |::|||  ||||   :
Db         90 VVPPAATLVLTGHWPLGATGCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALV 149

Qy        119 RPRYAGLLLGCAWGQSLAFSGAALGCSW--LGYSSAFASCSLRLPPEPERPRFAAFTATL 176
              | |   :  | | | | | :  | :|  :   |                 ||  || : :
Db        150 TKRRARAAVVLVWVVSAAVSFAPIMSQWWRVGADAEAQRCH-------SNPRCCAFASNM 202

Qy        177 ------HAVGFVLPLAVLCLTSLQVHRVARRHCQRMD------TVTMKALALLADLHPSV 224
                   :| | ||| |:     :|  || |  : :                : ||   | |:
Db        203 PYVLLSSSVSFYLPLLVMLFVYARVFVVATRQLRLLRWELGRFPPEESSPALSRSLAPAP 262

Qy        225 RQRCLIQQ------------KRRRHRATRKIGIAIATFLICFAPYVMTRL------AEL 265
                |  :               | |||   :|: : || :|: |: : :        |
Db        263 AGTCAPPEGVPACCRRPARLLPLREHRALCTLGLIMGTFTLCWLPFFLANVLRALGGPSL 322

Qy        266 VP---FVTVNAQWGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTP-- 317
              ||   |: :|  |       | |: :  :|  |       |  ||::|     | | | |
Db        323 VPDPAFLALN--W------LGYANSAFNPLIYCRSPDFRSAFRRLLCHCGGR-LPREPCA 373

Qy        318 --RPAST 322
                ||||:
Db        374 ADRPASS 380


RESULT 14
OX1R_HUMAN
ID   OX1R_HUMAN     STANDARD;      PRT;   425 AA.
AC   O43613;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Orexin receptor type 1 (Ox1r) (Hypocretin receptor type 1).
GN   HCRTR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98150861; PubMed=9491897;
RA   Sakurai T., Amemiya A., Ishii M., Matsuzaki I., Chemelli R.M.,
RA   Tanaka H., Williams S.C., Richardson J.A., Kozlowski G.P., Wilson S.,
RA   Arch J.R.S., Buckingham R.E., Haynes A.C., Carr S.A., Annan R.S.,
RA   McNulty D.E., Liu W.-S., Terrett J.A., Elshourbagy N.A., Bergsma D.J.,
RA   Yanagisawa M.;
RT   "Orexins and orexin receptors: a family of hypothalamic neuropeptides
RT   and G protein-coupled receptors that regulate feeding behavior.";
RL   Cell 92:573-585(1998).
RN   [2]
RP   REVIEW.
RX   MEDLINE=21237974; PubMed=11340621;
RA   Hungs M., Mignot E.;
RT   "Hypocretin/orexin, sleep and narcolepsy.";
RL   Bioessays 23:397-408(2001).
RN   [3]
RP   REVIEW.
RX   MEDLINE=21178476; PubMed=11283317;
RA   Willie J.T., Chemelli R.M., Sinton C.M., Yanagisawa M.;
RT   "To eat or to sleep? Orexin in the regulation of feeding and
RT   wakefulness.";
RL   Annu. Rev. Neurosci. 24:429-458(2001).
CC   -!- FUNCTION: Moderately selective excitatory receptor for orexin-A
CC       and, with a lower affinity, for orexin-B neuropeptide. Seems to be
CC       exclusively coupled to the G(q) subclass of heteromeric G
CC       proteins, which activates the phospholipase C mediated signaling
CC       cascade (By similarity).
```

```
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF041243; AAC39601.1; -.
DR   Genew; HGNC:4848; HCRTR1.
DR   MIM; 602392; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004930; F:G-protein coupled receptor activity; TAS.
DR   GO; GO:0007631; P:feeding behavior; TAS.
DR   GO; GO:0007218; P:neuropeptide signaling pathway; TAS.
DR   GO; GO:0007268; P:synaptic transmission; TAS.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR004059; Orexin_receptor1.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01521; OREXIN1R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane.
FT   DOMAIN       1     46       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    47     67       1 (POTENTIAL).
FT   DOMAIN      68     80       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    81    102       2 (POTENTIAL).
FT   DOMAIN     103    119       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   120    142       3 (POTENTIAL).
FT   DOMAIN     143    164       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   165    185       4 (POTENTIAL).
FT   DOMAIN     186    216       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   217    239       5 (POTENTIAL).
FT   DOMAIN     240    298       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   299    321       6 (POTENTIAL).
FT   DOMAIN     322    336       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   337    360       7 (POTENTIAL).
FT   DOMAIN     361    425       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD   194    194       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   425 AA;  47521 MW;  1634083DE10CA092 CRC64;

  Query Match           12.4%;  Score 229.5;  DB 1;  Length 425;
  Best Local Similarity  25.3%;  Pred. No. 4.5e-10;
  Matches   94;  Conservative   55;  Mismatches  154;  Indels   69;  Gaps   13;

Qy          5 EALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTL 64
              | :|    | |  |||: | || |    :  :||  :   :|||||  :|: |: :| :|
Db         46 EWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRT-VTNYFIVNLSLADVLVTAICLPASL 104

Qy         65 LGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAG 124
              |  :          |:|| :|     | | |::: :: |:| |:  || :    | | |
Db        105 LVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTAR-RARG 163

Qy        125 LLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAA----------FTA 174
               :||  |  |||          :  :|   ||  ||     | |  :          :
Db        164 SILG-IWAVSLAI---------MVPQAAVMECSSVLPELANRTRLFSVCDERWADDLYPK 213

Qy        175 TLHAVGFVLP-LAVLCLTSLQVHRVARRHCQRMDTVTMKALA--------LLADL----- 220
               |:  |::  || | | ::   :: |:    |    |  ||          | ||
Db        214 IYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLS 273

Qy        221 -HPSVRQRCL---IQQKRRRHRATRKIGIAIATFLICFAP----------YVMTRLA--- 263
               | | |     ::| | | :  :  :  :  :  | :|: |          : | | |
Db        274 GEPQPRGRAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDR 333

Qy        264 ELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRPFRQVL----------AGMVHRLL 313
               | |       : |      | |: : |:|  |: |   ||:            |    |
Db        334 EAVYACFTFSHW------LVYANSAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLK 387

Qy        314 KRTPRPASTHDS 325
               :|| :::| |
Db        388 APSPRSSASHKS 399


RESULT 15
SSR5_MOUSE
ID   SSR5_MOUSE     STANDARD;      PRT;   362 AA.
AC   O08858; O08998;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Somatostatin receptor type 5 (SS5R).
GN   SSTR5 OR SMSTR5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ; TISSUE=Liver;
RX   MEDLINE=97444289; PubMed=9300821;
RA   Lublin A.L., Diehl N.L., Hochgeschwender U.;
RT   "Isolation and characterization of the gene encoding the type 5 mouse
RT   (Mus musculus) somatostatin receptor (msst5).";
RL   Gene 195:63-66(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RA   Moldovan S., Demayo F., Brunicardi F.C.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Liver;
RA   Gordon D.F., Woodmansee W.W., Wood W.M., Knauf H., James R.A.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RA   Baumeister H., Roosterman D., Schafer J., Kreuzer O., Meyerhof W.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Receptor for somatostatin-28. The activity of this
CC       receptor is mediated by G proteins which inhibit adenylyl cyclase.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U82697; AAC53353.1; -.
DR   EMBL; AF004740; AAB61418.1; -.
DR   EMBL; AF030441; AAB86492.1; -.
DR   EMBL; AF035777; AAB88302.1; ALT_INIT.
DR   HSSP; P02699; 1F88.
DR   MGD; MGI:894282; Sstr5.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Transmembrane; Glycoprotein;
KW   Multigene family; Lipoprotein; Palmitate.
FT   DOMAIN       1     35       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    36     63       1 (POTENTIAL).
FT   DOMAIN      64     73       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    74     99       2 (POTENTIAL).
FT   DOMAIN     100    110       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   111    132       3 (POTENTIAL).
```

```
FT   DOMAIN      133    154       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    155    175       4 (POTENTIAL).
FT   DOMAIN      176    195       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    196    220       5 (POTENTIAL).
FT   DOMAIN      221    246       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    247    272       6 (POTENTIAL).
FT   DOMAIN      273    282       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    283    307       7 (POTENTIAL).
FT   DOMAIN      308    362       CYTOPLASMIC (POTENTIAL).
FT   CARBOHYD     13     13       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     23     23       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    185    185       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    109    184       BY SIMILARITY.
FT   CONFLICT     99     99       V -> VV (IN REF. 1).
FT   CONFLICT    303    305       YGF -> LWL (IN REF. 2).
SQ   SEQUENCE   362 AA;  39948 MW;  AA091DDD570FDFAB CRC64;

  Query Match             12.3%;  Score 229;  DB 1;  Length 362;
  Best Local Similarity   25.7%;  Pred. No. 4.2e-10;
  Matches   95;  Conservative   53;  Mismatches  162;  Indels   60;  Gaps   13;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              ||     |:  | ::|  | |  | ||:      |:::|  : | ::||::  :|   | :
Db         33 MGARAVLVPVLYLLVCTVGLGGNTLVIYVVLRYAKMKT-VTNVYILNLAVADVLF-MLGL 90

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              ||          |     |:::  ||      :   :| |::|||  ||| |   ||
Db         91 PFLATQNAVSYWPFGSFLCRLVMTLDGINQFTSIFCLMVMSVDRYLAVVHPLRSARWRRP 150

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              | | |      |  ||  |   |   :      : :|:|   |||     |||      :|
Db        151 RVAKLASAAVWVFSLLMSLPLL--VFADVQEGWGTCNLSW-PEPVGLWGAAFITYTSVLG 207

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |  || |:||             |  :  | :||  :      | |:|      |   : |
Db        208 FFGPLLVICL-----------CYLLIVVKVKAAGMRVG---SSRRR------RSERKVT 246

Qy        241 RKIGIAIATFLICFAPYVMTRLAEL---VPFVTVNAQWGILSKCLTYSKAVADPFTYSLL 297
              | : : :  |: |: |: :  :  |   :|    :|        |:|: : |:| | |
Db        247 RMVVVVVLVFVGCWLPFFIVNIVNLAFTLPEEPTSAGLYFFVVVLSYANSCANPLLYGFL 306

Qy        298 ----RRPFRQVLAGMVHRLLKR----------TPRPASTHDSSLDVAGMVHQLLKRTPRP 343
                 |: ||: |      |:|          |||       | :|     |  |
Db        307 SDNFRQSFRKALC------LRRGYGVEDADAIEPRP--------DKSGRPQTTL---PTR 349

Qy        344 ASTHNGSVDT 353
              :   || : |
Db        350 SCEANGLMQT 359


Search completed: May 18, 2004, 12:07:09
Job time : 20 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 18, 2004, 12:03:02 ; Search time 46 Seconds
(without alignments)
2489.850 Million cell updates/sec

Title: US-10-049-569-2
Perfect score: 1858
Sequence: 1 MGPGEALLAGLLVMVLAVAL.....ASTHNGSVDTENDSCLQQTH 363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	250.5	13.5	438	11	Q63004	Q63004 rattus norv
2	246.5	13.3	390	13	Q8QGQ4	Q8qgq4 carassius a
3	238.5	12.8	380	13	Q9DGQ6	Q9dgq6 carassius a
4	237	12.8	370	13	Q8UWL5	Q8uwl5 fugu rubrip
5	237	12.8	399	5	Q86ME4	Q86me4 caenorhabdi
6	237	12.8	402	5	Q86ME5	Q86me5 caenorhabdi
7	233	12.5	478	11	Q8BLD9	Q8bld9 mus musculu
8	231.5	12.5	353	13	Q9YHY3	Q9yhy3 lampetra fl
9	230.5	12.4	425	4	Q9HBV6	Q9hbv6 homo sapien
10	230	12.4	367	6	Q8MI04	Q8mi04 ovis aries
11	229.5	12.4	340	13	Q9YHY2	Q9yhy2 lampetra fl
12	229	12.3	385	11	Q9JK40	Q9jk40 mus musculu
13	228.5	12.3	437	13	O42316	O42316 cyprinus ca
14	227.5	12.2	367	13	Q9PVG0	Q9pvg0 carassius a
15	227.5	12.2	391	5	O96716	O96716 branchiosto
16	227.5	12.2	477	4	Q8NEQ8	Q8neq8 homo sapien
17	227	12.2	380	13	O42317	O42317 cyprinus ca
18	226.5	12.2	358	11	Q9QX37	Q9qx37 mus musculu
19	226.5	12.2	397	11	Q9D282	Q9d282 mus musculu
20	226	12.2	390	13	Q8AXM7	Q8axm7 carassius a
21	224	12.1	367	13	Q9PVF9	Q9pvf9 carassius a
22	223.5	12.0	371	11	Q8BKB0	Q8bkb0 mus musculu
23	223	12.0	387	4	Q96KI0	Q96ki0 homo sapien
24	222.5	12.0	385	11	Q8BQ97	Q8bq97 mus musculu
25	222.5	12.0	447	6	Q8SPH2	Q8sph2 sus scrofa
26	222	11.9	388	4	Q8IXH9	Q8ixh9 homo sapien
27	221.5	11.9	446	13	O42315	O42315 cyprinus ca
28	221	11.9	445	13	Q98841	Q98841 anguilla an
29	220.5	11.9	388	5	Q8MTW6	Q8mtw6 spisula sol
30	220	11.8	397	4	Q7Z5R9	Q7z5r9 homo sapien
31	219.5	11.8	315	6	Q9GKP7	Q9gkp7 sus scrofa
32	219	11.8	470	11	P97842	P97842 rattus norv
33	218	11.7	394	5	Q9NHF3	Q9nhf3 aplysia cal
34	217	11.7	452	13	Q7T2S8	Q7t2s8 carassius a
35	216.5	11.7	346	6	Q95KS6	Q95ks6 ovis aries
36	216	11.6	460	5	Q86ME6	Q86me6 caenorhabdi
37	215	11.6	331	13	Q801F3	Q801f3 fugu rubrip
38	214	11.5	385	4	Q86YF2	Q86yf2 homo sapien
39	214	11.5	559	11	Q9QW71	Q9qw71 rattus sp.
40	213.5	11.5	518	6	Q9MYI8	Q9myi8 oryctolagus
41	212	11.4	394	5	Q9NJS6	Q9njs6 aplysia kur
42	211.5	11.4	352	13	Q9YHY4	Q9yhy4 lampetra fl
43	211	11.4	445	13	Q98842	Q98842 anguilla an
44	210.5	11.3	459	11	Q8CGI5	Q8cgi5 mus musculu
45	210.5	11.3	515	11	Q9DBL0	Q9dbl0 mus musculu

ALIGNMENTS

RESULT 1
Q63004
ID Q63004 PRELIMINARY; PRT; 438 AA.
AC Q63004;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 5-HT6 serotonin receptor.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Striatum;
RX MEDLINE=96102917; PubMed=8522988;
RA Kohen R., Metcalf M.A., Khan N., Druck T., Huebner K., Lachowicz J.E.,
RA Meltzer H.Y., Sibley D.R., Roth B.L., Hamblin M.W.;
RT "Cloning, characterization, and chromosomal localization of a human 5-
RT HT6 serotonin receptor.";
RL J. Neurochem. 66:47-56(1996).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; L41146; AAA92633.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 438 AA; 46786 MW; DB90403892F691A9 CRC64;

Query Match 13.5%; Score 250.5; DB 11; Length 438;
Best Local Similarity 24.7%; Pred. No. 7e-13;
Matches 98; Conservative 69; Mismatches 176; Indels 53; Gaps 14;

```
Qy          2 GPGEA--------LLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHL 53
Db         16 GPGPPPAPGGSGWVAAALCVVIVLTAAANSLLIVLICTQPAVRNT--SNFFLVSLFTSDL 73

Qy         54 LLAALDMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLR 113
Db         74 MVGLVVMPPAMLNALYGRWVLARGLCLLWTAFDVMCCSASILNLCLISLDRYLLILSPLR 133

Qy        114 YAGRL-RPRYAGLLLGCAWGQSLA----FSGAALGCSWLGYSSAFASCSLRLPPEPERPR 168
Db        134 YKLRMTAPRALALILG-AW--SLAALASFLPLLLGWHELGKART---------PAPGQCR 181

Qy        169 FAA---FTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVT-------MKALALLA 218
Db        182 LLASLPFVLVASGVTFFLPSGAICFTYCRILLAARKQAVQVASLTTGTAGQALETLQVPR 241

Qy        219 DLHPSVR----QRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQ 274
Db        242 TPRPGMESADSRRLATKHSRKALKASLTLGILLGMFFVTWLPFFVANIAQAVCDCISPGL 301

Qy        275 WGILSKCLTYSKAVADPFTYSLLRRPFRQVLAGMV--------HRLLKRTPRPASTHDSS 326
Db        302 FDVLT-WLGYCNSTMNPIIYPLFMRDFKRALGRFLPCVHCPPEHRASPASPSMWTSHSGA 360

Qy        327 LDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQQT 362
Db        361 R--PGLSLQQVLPLPLPPNSDSDSA-SGGTSGLQLT 393
```

RESULT 2
Q8QGQ4
ID   Q8QGQ4   PRELIMINARY;   PRT;   390 AA.
AC   Q8QGQ4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Type five-like somatostatin receptor.
OS   Carassius auratus (Goldfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Carassius.
OX   NCBI_TaxID=7957;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lin X., Peter R.E.;
RT   "Cloning and expression of a type five-like somatostatin receptor in
RT   goldfish.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF252879; AAM18805.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001092; HLH_basic.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00038; HLH_1; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   390 AA; 44136 MW; EC12119A4B6CF9A8 CRC64;

Query Match 13.3%; Score 246.5; DB 13; Length 390;
Best Local Similarity 23.9%; Pred. No. 1.3e-12;
Matches 79; Conservative 59; Mismatches 122; Indels 71; Gaps 10;

```
Qy          7 LLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLG 66
Db         42 ILAVIYLVVFVVGLTGNSLAIFVVLRYTKMKT-ATNMYILNLAVADELY-ILGLPFLTAH 99

Qy         67 VMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLL 126
Db        100 NMLGYWPFGNFLCRILMWTDSISQFTSTFCLTVMSIDRYMAVVHPIRSARWRRPRVAKVI 159

Qy        127 LGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLP-----------------PEPERPR 168
Db        160 NSMVWA--------------------LSCLLTLPVIIYCDVQPELNTCNLSWPEPRDVW 198

Qy        169 FAAFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRC 228
Db        199 STAFILYTAMLGFFFPLMVICL------------CYLLIVIKVKSASARAGLS------- 239

Qy        229 LIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVNA-QWGI--LSKCLTYS 285
Db        240 --KRGRSEKKVTRMVVIIVVVFVLCWLPFFIINILNLISTLPENSLMTGIYFLTVILTYV 297

Qy        286 KAVADPFTYSLL----RRPFRQVLAGMVHRL 312
Db        298 NSCANPLLYSFLSDNFKRSFQQVLC--IHKV 326
```

RESULT 3
Q9DGQ6
ID   Q9DGQ6   PRELIMINARY;   PRT;   380 AA.
AC   Q9DGQ6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Somatostatin receptor type two.
GN   SST2.
OS   Carassius auratus (Goldfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Carassius.
OX   NCBI_TaxID=7957;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20453044; PubMed=10996426;
RA   Lin X., Janovick J.A., Cardenas R., Conn P.M., Peter R.E.;
RT   "Molecular cloning and expression of a type-two somatostatin receptor
RT   in goldfish brain and pituitary.";
RL   Mol. Cell. Endocrinol. 166:75-87(2000).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF139597; AAF98367.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   380 AA; 43146 MW; CE897FF7537CFA30 CRC64;

Query Match 12.8%; Score 238.5; DB 13; Length 380;
Best Local Similarity 24.1%; Pred. No. 6.1e-12;
Matches 84; Conservative 63; Mismatches 163; Indels 39; Gaps 9;

```
Qy         14 MVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGVMRGRTP 73
Db         60 VVCAVGLCGNALVMYVILRYAKMKT-VTNIYILNLAVAD-VLCMLSLPFIAIQLSLLHWP 117

Qy         74 SAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLLLGCAWGQ 133
Db        118 FGSAICRVVLTADSMNQFTSIFFLTVMSFDRYLAVVHPIKSTKWRKPRMAKSISLAMWVI 177

Qy        134 SLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVGFVLPLAVLCLTSL 193
Db        178 SLLVNLPIMIYSGVNVKKNEARTCTMLWPEPQNTYYTVFIFYTFFMGFFLPLIVICM--- 234
```

```
Qy        194 QVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATFLIC 253
Db        235 ---------CYLLIVIKVKS---------SGMRVCSSKRKRSERKVTRMVSIVVVVFVLC 276

Qy        254 FAPYV---MTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLL----RRPFRQVLA 306
Db        277 WLPFYVFNVTSVTGTVPTTPVLKSTFDFVVVLGYANSCANPILYAFLSDNFKKSFQNVLC 336

Qy        307 GMVHRLLKRTP--RPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDT 353
Db        337 ------LKRVGGLDEIERSDSRQDRTRMVNDVMSET-HNAALLNGDLQT 378


RESULT 4
Q8UWL5
ID   Q8UWL5      PRELIMINARY;      PRT;   370 AA.
AC   Q8UWL5;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Somatostatin receptor 2.
GN   SSTR2.
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21564205; PubMed=11707075;
RA   Bagheri-Fam S., Ferraz C., Demaille J., Scherer G., Pfeifer D.;
RT   "Comparative Genomics of the SOX9 Region in Human and Fugu rubripes:
RT   Conservation of Short Regulatory Sequence Elements within Large
RT   Intergenic Regions.";
RL   Genomics 78:73-82(2001).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF329945; AAL32173.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   370 AA;  41364 MW;  420BB12F204946B6 CRC64;

  Query Match             12.8%;  Score 237;  DB 13;  Length 370;
  Best Local Similarity   22.9%;  Pred. No. 7.9e-12;
  Matches   86;  Conservative   61;  Mismatches  133;  Indels   96;  Gaps   14;

Qy         14 MVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGVMRGRTP 73
Db         53 LVCAVGLCGNALVIYVILRYAKMKT-VTNIYILNLAVADVLF-MLGLPFIAIQLALVHWP 110

Qy         74 SAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLLLGCAWGQ 133
Db        111 FGPVLCRVVMTVDSLNQFTSIFCLMVMSIDRYLAVVHPIKSTKWRKPRVAKTINVAVWGA 170

Qy        134 SLA-------FSGAAL---GCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVGFVL 183
Db        171 SLVVNLPIVIYSGIITKQDGC----------FCTI-VWPEPEEAYYTAFMIYTFILGFFL 219

Qy        184 PLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRATRKI 243
Db        220 PLLVISL------------CYVFIIVKVKSSGIRVGSS---------KRKRSERKVTRMV 258

Qy        244 GIAIATFLICFAPYVMTRLAELV------------PFVTVNAQWGILSKCLTYSKAVAD 290
Db        259 SIVVAVFVLCWLPFYVFNVTSVTGSISATHVLRSTFAFVVV----------LGYANSCAN 308

Qy        291 PFTYSLLRRPFR---------QVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTP 341
Db        309 PILYAFLSENFKKSFQNVLCLQKVGGL---------DEAERSDSRQDKSRMIND------ 353

Qy        342 RPASTH----NGSVDT 353
Db        354 -PTETRSTLLNGDLQT 368


RESULT 5
Q86ME4
ID   Q86ME4      PRELIMINARY;      PRT;   399 AA.
AC   Q86ME4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein F15A8.5.
GN   F15A8.5.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Waterston R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Miller N.;
RT   "The sequence of C. elegans cosmid F15A8.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U97549; AAO91738.1; -.
DR   WormPep; F15A8.5c; CE33617.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000412; ABC_transpt2.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00890; ABC2_MEMBRANE; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   399 AA;  45526 MW;  40629EF6E074D048 CRC64;

  Query Match             12.8%;  Score 237;  DB 5;  Length 399;
  Best Local Similarity   23.2%;  Pred. No. 8.6e-12;
  Matches   82;  Conservative   61;  Mismatches  142;  Indels   68;  Gaps   11;

Qy          8 LAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGV 67
Db          8 LLGLFSVLIILALFGNLLVCAAILWDRSLRKQPENLFLVSLAVSDLLVSVLVMLFAAVND 67

Qy         68 MRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFP---LRYAGRLRPRYAG 124
Db         68 ILGYWPFGQFYCQFWISFDITTCTASILNLCAISLDRYWHISRPMVYIRYCNRRRINYVI 127
```

```
Qy        125 LLLGCAWGQSLAFSGAALGCSWLGYSS--------AFASCSLRLPPEPERPRFAAFTATL 176
Db        128 VLV---WLISAGIGAAPLG---FGFGSKVTINNLTGLPVCEMRLP--------LPYAIGS 173

Qy        177 HAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALL---------ADLHPSVRQR 227
Db        174 SMVSFFLPAMVMVILYTKLYLYARKHVRSIKTQLQQATSFLIMQLASEKIREIRTSIFSK 233

Qy        228 ----CLIQQKRRR-----------------HRATRKIGIAIATFLICFAPYVMTR---- 261
Db        234 LNFLCPTRFKNQRSPQDPHTPAAHNRSNISDQKARLTLGVIMGTFLVCWLPFFTVNILRA 293

Qy        262 -LAELVPFVTVNA-QWGILSKCLTYSKAVADPFTYSLLRRPFRQVLAGMVHRL 312
Db        294 WLPEIFSSKTIMAVTW------LGYANSSANPLIYSIFNRDFRRAFKKIIVRV 340


RESULT 6
Q86ME5
ID   Q86ME5      PRELIMINARY;      PRT;   402 AA.
AC   Q86ME5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein F15A8.5.
GN   F15A8.5.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RA   Waterston R.;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology. The C. elegans Sequencing Consortium.";
RL   Science 282:2012-2018(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Miller N.;
RT   "The sequence of C. elegans cosmid F15A8.";
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Waterston R.;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U97549; AAO91737.1; -.
DR   WormPep; F15A8.5b; CE33616.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000412; ABC_transpt2.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00890; ABC2_MEMBRANE; 1.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   402 AA;  45856 MW;  06ABE42C0970CBF5 CRC64;

  Query Match            12.8%;  Score 237;  DB 5;  Length 402;
  Best Local Similarity  23.2%;  Pred. No. 8.7e-12;
  Matches   82;  Conservative   61;  Mismatches  142;  Indels   68;  Gaps   11;

Qy          8 LAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGV 67
Db          8 LLGLFSVLIILALFGNLLVCAAILWDRSLRKQPENLFLVSLAVSDLLVSVLVMLFAAVND 67

Qy         68 MRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFP---LRYAGRLRPRYAG 124
Db         68 ILGYWPFGQFYCQFWISFDITTCTASILNLCAISLDRYWHISRPMVYIRYCNRRRINYVI 127

Qy        125 LLLGCAWGQSLAFSGAALGCSWLGYSS--------AFASCSLRLPPEPERPRFAAFTATL 176
Db        128 VLV---WLISAGIGAAPLG---FGFGSKVTINNLTGLPVCEMRLP--------LPYAIGS 173

Qy        177 HAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALL---------ADLHPSVRQR 227
Db        174 SMVSFFLPAMVMVILYTKLYLYARKHVRSIKTQLQQATSFLIMQLASEKIREIRTSIFSK 233

Qy        228 ----CLIQQKRRR-----------------HRATRKIGIAIATFLICFAPYVMTR---- 261
Db        234 LNFLCPTRFKNQRSPQDPHTPAAHNRSNISDQKARLTLGVIMGTFLVCWLPFFTVNILRA 293

Qy        262 -LAELVPFVTVNA-QWGILSKCLTYSKAVADPFTYSLLRRPFRQVLAGMVHRL 312
Db        294 WLPEIFSSKTIMAVTW------LGYANSSANPLIYSIFNRDFRRAFKKIIVRV 340


RESULT 7
Q8BLD9
ID   Q8BLD9      PRELIMINARY;      PRT;   478 AA.
AC   Q8BLD9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   D.
GN   DRD5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK045456; BAC32378.1; -.
DR   PIR; PT0721; PT0721.
DR   MGD; MGI:94927; Drd5.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
SQ   SEQUENCE   478 AA;  53076 MW;  D39165BC558131DF CRC64;

  Query Match            12.5%;  Score 233;  DB 11;  Length 478;
  Best Local Similarity  22.9%;  Pred. No. 2.3e-11;
  Matches   82;  Conservative   76;  Mismatches  146;  Indels   54;  Gaps   13;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
Db         34 LGPAQVVTAGLLTLLIVWTLLGNVLVCAAIVRSRHLRAKMTNIFIVSLAVSDLFVALLVM 93

Qy         61 PFTLLGVMRGRTPSAPGA-CQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLR 119
Db         94 PWKAVAEVAGYWPF--GAFCDIWVAFDIMCSTASILNLCIISVDRYWAISRPFRYERKMT 151

Qy        120 PRYAGLLLGCAWGQSLAFSGAALGCSW----LGYSSAFASCSLRLPPE------------ 163
```

```
Db        152 QRVALVMVALAWTLSILISFIPVQLNWHRDKAGSQGREGLLSNETPWEEGWELDGRTENC 211

Qy        164 -PERPRFAAFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHP 222
                |  | :::|  :  | :|:|:: :|  :::|:|:   :|     : :|   |:
Db        212 DSSLNRTYAISSSL--ISFYIPVAIMIVTYTRIYRIAQVQIRR-----ISSLERAAEHAQ 264

Qy        223 SVRQR-------CLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQW 275
              | |          |    ::  :  : : : :  |: |: |: :  | :||| :
Db        265 SCRSRGACEPDPSLRASIKKETKVFKTLSVIMGVFVCCWLPFFI--LNCMVPFCSSGDAQ 322

Qy        276 GILS--KCLT-----------YSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTP 317
              |  :   |::          :: :  :|  |:      |: | |:| |  | |  |||
Db        323 GPRTGFPCVSETTFDIFVWFGWANSSLNPIIYAFNADFRKVFAQLL-GCSH-LCFRTP 378


RESULT 8
Q9YHY3
ID   Q9YHY3      PRELIMINARY;      PRT;   353 AA.
AC   Q9YHY3;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative odorant receptor LOR4.
OS   Lampetra fluviatilis (River lamprey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Lampetra.
OX   NCBI_TaxID=7748;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Olfactory organ;
RA   Berghard A., Dryer L.;
RT   "A novel family of ancient vertebrate odorant receptors.";
RL   J. Neurobiol. 0:0-0(1998).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF069547; AAC82382.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001092; HLH_basic.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR   PROSITE; PS00038; HLH_1; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   353 AA;  39901 MW;  D89FB9D6FC827180 CRC64;

  Query Match           12.5%;  Score 231.5;  DB 13;  Length 353;
  Best Local Similarity   22.0%;  Pred. No. 2.2e-11;
  Matches   76;  Conservative   88;  Mismatches  147;  Indels   35;  Gaps   10;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              |    | |: :  ::||  :| | |::   ||  :|:|| : :: ::|::  ||:    |
Db         28 MSERTAALSAITTLILA-TVLGNLLIITSIAYFRQLQTR-TNIMALSLAVADLLVGLTVM 85

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              |::::  :           |  :  |||  | ::: : :  :: |:::|:  ||||  |:
Db         86 PYSMMKAVYKCWFYGQFFCNLQYFLDYMLTNSSIMHLGCIAYDRYVAICDPLRYPQRVTN 145

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
                :|  :|     ||    |    : ::  ::  :: :   | :  |   |  :  ||
Db        146 HTVVTMLLISWFGPALFSSPIL----VNFNPEWSRGNIEIISCPNQCLFFVSTWLVSVVG 201

Qy        181 ---FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQR------CLIQ 231
                 :|| |  :     ::: ||||  ::: :|:::       :|   :|:        |:
Db        202 VCPYVLSLLTMSYVYARIYIVARRQGRQISSVSLQ-------VHAQQQQQQQQAEPTNIR 254

Qy        232 QK----RRRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKA 287
              ||    :| | | : :|  |  :|: : |: | |     |    ::    ::  : |  :
Db        255 QKWTAMKREHSAAKTLGSIIGVYLLSWLPFYMVVL--FFPNFQNSSAAVRITTWIGYISS 312

Qy        288 VADPFTYSLLRRPFRQVLAGMVH-RLLKRTPRPASTHDSSLDVAGM 332
                :|  |: | ||||     ::  ::|      :||   ::|::|:
Db        313 AINPVLYATLNRPFRSAFVAVISCKVL------SSTRARTMDLSGV 352


RESULT 9
Q9HBV6
ID   Q9HBV6      PRELIMINARY;      PRT;   425 AA.
AC   Q9HBV6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypocretin receptor-1.
GN   HCRTR1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20429525; PubMed=10973318;
RA   Peyron C., Faraco J., Rogers W., Ripley B., Overeem S., Charnay Y.,
RA   Nevsimalova S., Aldrich M., Reynolds D., Albin R., Li R., Hungs M.,
RA   Pedrazzoli M., Padigaru M., Kucherlapati M., Fan J., Maki R.,
RA   Lammers G.J., Bouras C., Kucherlapati R., Nishino S., Mignot E.;
RT   "A mutation in a case of early onset narcolepsy and a generalized
RT   absence of hypocretin peptides in human narcoleptic brains.";
RL   Nat. Med. 6:991-997(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Yeager M., Welch R., Haque K., Bergen A.;
RT   "Genomic sequence of the hypocretin (orexin) receptor 1 (HCRTR1).";
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21580342; PubMed=11723285;
RA   Olafsdottir B.R., Rye D.B., Scammell T.E., Matheson J.K.,
RA   Stefansson K., Gulcher J.R.;
RT   "Polymorphisms in hypocretin/orexin pathway genes and narcolepsy.";
RL   Neurology 57:1896-1899(2001).
DR   EMBL; AF202084; AAG28020.1; -.
DR   EMBL; AF202078; AAG28020.1; JOINED.
DR   EMBL; AF202079; AAG28020.1; JOINED.
DR   EMBL; AF202080; AAG28020.1; JOINED.
DR   EMBL; AF202081; AAG28020.1; JOINED.
DR   EMBL; AF202082; AAG28020.1; JOINED.
DR   EMBL; AF202083; AAG28020.1; JOINED.
DR   EMBL; AY070269; AAL50221.1; -.
DR   EMBL; AY062030; AAL47214.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0016499; F:orexin receptor activity; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR004059; Orexin_receptor1.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PRINTS; PR01521; OREXIN1R.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   Receptor.
SQ   SEQUENCE   425 AA;  47535 MW;  B650B37F3A2CA096 CRC64;

  Query Match           12.4%;  Score 230.5;  DB 4;  Length 425;
  Best Local Similarity   25.3%;  Pred. No. 3.3e-11;
  Matches   94;  Conservative   55;  Mismatches  154;  Indels   69;  Gaps   13;

Qy          5 EALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTL 64
              | :|    | |  |||: | || |    :  :|| :   :|||||  :|: |: :| :|
Db         46 EWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRT-VTNYFIVNLSLADVLVTAICLPASL 104
```

```
Qy          65 LGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAG 124
Db         105 LVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTAR-RARG 163

Qy         125 LLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAA----------FTA 174
Db         164 SILG-IWAVSLAI---------MVPQAAVMECSSVLPELANRTRLFSVCDERWADDLYPK 213

Qy         175 TLHAVGFVLP-LAVLCLTSLQVHRVARRHCQRMDTVTMKALA--------LLADL----- 220
Db         214 IYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLS 273

Qy         221 -HPSVRQRCL---IQQKRRRHRATRKIGIAIATFLICFAP----------YVMTRLA--- 263
Db         274 GEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDR 333

Qy         264 ELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRPFRQVL----------AGMVHRLL 313
Db         334 EAVYACFTFSHW------LVYANSAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLK 387

Qy         314 KRTPRPASTHDS 325
Db         388 APSPRSSASHKS 399
```

RESULT 10
Q8MI04
ID Q8MI04 PRELIMINARY; PRT; 367 AA.
AC Q8MI04;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Somatostatin receptor subtype 5.
GN SST5.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RA Debus N., Dutour A., Boudouresque F., Vuaroqueaux V., Oliver C.,
RA Ouafik L.;
RT "Molecular cloning and tissue distribution of the ovine somatostatin
RT receptor subtype 5: osst5.";
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RA Debus N.;
RT "Regulation of GH secretion in sheep: Role of somatostatin and its
RT receptors.";
RL Thesis (1999),
RL Department of Intercellular communications in Endocrinology,
RL University of Aix-Marseilles II, Marseilles, France.
DR EMBL; AJ441116; CAD29615.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT CHAIN 2 367 SOMATOSTATIN RECEPTOR SUBTYPE 5.
SQ SEQUENCE 367 AA; 40269 MW; FA9CAB7AA8BBD764 CRC64;

Query Match 12.4%; Score 230; DB 6; Length 367;
Best Local Similarity 26.2%; Pred. No. 3.1e-11;
Matches 79; Conservative 51; Mismatches 138; Indels 34; Gaps 8;

```
Qy          11 LLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGVMRG 70
Db          46 LYLLVCAVGLGGNALVIYVVLRHAKMKT-VTNIYILNLAVADVLL-MLGLPFVATQNAIS 103

Qy          71 RTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLLLGCA 130
Db         104 YWPFGPVLCRLVMTLDGINQFTSIFCLTVMSVDRYLAVVHPIRSARWRRPRVAKLASAAV 163

Qy         131 WGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVGFVLPLAVLCL 190
Db         164 WAFSLVMSLPLV--VFADIQEGWNTCNLTW-PEPVGLWGAVFIIYTSVLGFFGPLLVICL 220

Qy         191 TSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATF 250
Db         221 -----------CYVLIVVKLKASGV----------RVGSTRRRSERKVTRMVVVVVLVF 258

Qy         251 LICFAPYVMTRLAEL---VPFVTVNAQWGILSKCLTYSKAVADPFTYSL----LRRPFRQ 303
Db         259 AGCWLPFFIVNIVNLAFALPEEPASAGLYFFVVVLSYANSCANPCLYGFLSDNLRQSFRK 318

Qy         304 VL 305
Db         319 VL 320
```

RESULT 11
Q9YHY2
ID Q9YHY2 PRELIMINARY; PRT; 340 AA.
AC Q9YHY2;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative odorant receptor LOR14.
OS Lampetra fluviatilis (River lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lampetra.
OX NCBI_TaxID=7748;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory organ;
RA Berghard A., Dryer L.;
RT "A novel family of ancient vertebrate odorant receptors.";
RL J. Neurobiol. 0:0-0(1998).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF069548; AAC82383.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 340 AA; 37947 MW; E329864C73C18F25 CRC64;

Query Match 12.4%; Score 229.5; DB 13; Length 340;
Best Local Similarity 22.2%; Pred. No. 3.1e-11;
Matches 74; Conservative 65; Mismatches 141; Indels 53; Gaps 9;

```
Qy           2 GPGEALLAGLLVMV-----LAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLA 56
Db          23 GPTQVREVRLVLLIVFAFFITATILGNLLIITAIASFTKLQTHAN-FLALSLAVSDLLVG 81

Qy          57 ALDMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAG 116
Db          82 VFIMPLSMMKTVYDCWFYASVLCNAHYFLDYTLTTATILHITCIAYDRYVAICDPLRYAT 141

Qy         117 RLRPRYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATL 176
```

```
Db        142 RVTVRTIAALLVLCWLGAAIFSSPIL----LSFSPTLSRGNIERASCPDDCVFSVSFGIL 197

Qy        177 HAVG---FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQK 233
                :|   :   : |: |    ::|||||  :::                        ::
Db        198 VTIGMGPYFTAIHVVLLMYAVIYRVARRQSRKV----------------AADSSAVKNM 240

Qy        234 RRRHRATRKIGIAIATFLICFAPYVMTRLAE------LVPFVTVNAQWGILSKCLTYSKA 287
              :| | | : :|  | ||:: : || :  :          :|:    | |      : |  :
Db        241 KREHSAAKTLGAIIGTFMLSWLPYYVVAMMSFTVDGLFIPYRV--AMW------IGYCSS 292

Qy        288 VADPFTYSLLRRPFR---------QVLAGMVHR 311
                :|  |:   ||||         :| :|:| |
Db        293 AINPLLYASFNRPFRTAFRDIFRLRVFSGVVRR 325


RESULT 12
Q9JK40
ID   Q9JK40      PRELIMINARY;      PRT;   385 AA.
AC   Q9JK40;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Somatostatin subtype 5 receptor.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv;
RA   O'Carroll A.-M.;
RT   "Cloning, sequence and tissue distribution of the gene encoding a
RT   mouse somatostatin subtype 5 receptor.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF268067; AAF72547.1; -.
DR   HSSP; P02699; 1F88.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   385 AA;  42489 MW;  5EC1140BF09FCEB5 CRC64;

  Query Match             12.3%;  Score 229;  DB 11;  Length 385;
  Best Local Similarity   25.7%;  Pred. No. 3.9e-11;
  Matches   95;  Conservative   53;  Mismatches  162;  Indels   60;  Gaps   13;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              ||     |:  | ::|  | |  | ||:      |:::|  :  | ::||::  :|   | :
Db         56 MGARAVLVPVLYLLVCTVGLGGNTLVIYVVLRYAKMKT-VTNVYILNLAVADVLF-MLGL 113

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              ||          |      |:::  ||      :   :  :| |::|||  ||| |   ||
Db        114 PFLATQNAVSYWPFGSFLCRLVMTLDGINQFTSIFCLMVMSVDRYLAVVHPLRSARWRRP 173

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              | | |      |  ||  |   |  :      : :|:|   |||      |||       :|
Db        174 RVAKLASAAVWVFSLLMSLPLL--VFADVQEGWGTCNLSW-PEPVGLWGAAFITYTSVLG 230

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |  || |:||           | :  | :||  :      | |:|      |   : |
Db        231 FFGPLLVICL------------CYLLIVVKVKAAGMRVG---SSRRR------RSERKVT 269

Qy        241 RKIGIAIATFLICFAPYVMTRLAEL---VPFVTVNAQWGILSKCLTYSKAVADPFTYSLL 297
              | : : :  |: |: |: :  :  |   :|    :|        |:|: : |:|  |  |
Db        270 RMVVVVVLVFVGCWLPFFIVNIVNLAFTLPEEPTSAGLYFFVVVLSYANSCANPLLYGFL 329
```

```
Qy        298 ----RRPFRQVLAGMVHRLLKR----------TPRPASTHDSSLDVAGMVHQLLKRTPRP 343
                  |: ||: |       |:|          |||        | :|     |   |
Db        330 SDNFRQSFRKALC------LRRGYGVEDADAIEPRP--------DKSGRPQTTL---PTR 372

Qy        344 ASTHNGSVDT 353
                :   || : |
Db        373 SCEANGLMQT 382


RESULT 13
O42316
ID   O42316      PRELIMINARY;      PRT;   437 AA.
AC   O42316;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   D1A4 dopamine receptor.
OS   Cyprinus carpio (Common carp).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Cyprinus.
OX   NCBI_TaxID=7962;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Retina;
RX   MEDLINE=99044151; PubMed=9826915;
RA   Hirano J., Archer S.N., Djamgoz M.B.A.;
RT   "Dopamine receptor subtypes expressed in vertebrate (carp and eel)
RT   retinae: cloning, sequencing and comparison of five D1-like and three
RT   D2-like receptors.";
RL   Recept. Channels 5:387-404(1998).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; Y14627; CAA74971.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   437 AA;  48707 MW;  A44749475F2139CE CRC64;

  Query Match             12.3%;  Score 228.5;  DB 13;  Length 437;
  Best Local Similarity   22.3%;  Pred. No. 4.9e-11;
  Matches   74;  Conservative   79;  Mismatches  142;  Indels   37;  Gaps   13;

Qy          7 LLAG-LLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLL 65
              :| |  | :::   || | ||          ||:: :   :::|::  ||:| | ||:
Db         22 VLTGCFLSLLILTTLLGNTLVCAAVTKFRHLRSKVTNFFVISLAISDLLVAILVMPWKAA 81

Qy         66 GVMRGRTPSAPGA-CQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAG 124
                : |  |   || | |     |   :: : |:  :| |::  |:  | ||  :: |: |
Db         82 SEIVGFWPF--GAFCDVWVAFDIMCSTASILNLCVISVDRYWAISSPFRYERKMTPKVAF 139

Qy        125 LLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPP---EPERPRFAAFTATLHAVGF 181
              :::  ||  |:  |   :  :|     :        |||   :     |  | :::|  : |
Db        140 IMISLAWTLSILISFIPVQLNWHKAQAELNGTYGELPPDNCDSSLNRTYAISSSL--ISF 197

Qy        182 VLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQK-------R 234
              :|:|:: :|  :::|:|::  :|:  :  :|      : | |:     :: :       :
Db        198 YIPVAIMLVTYTRIYRIAQKQIRRISALE-RAAESAKNRHSSMGNNASMESESSFKMSFK 256

Qy        235 RRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLT----------- 283
              |  :  : : :  :|: |: |: :  |  :|||   |   |     |::
Db        257 RETKVLKTLSVIMGVFVCCWLPFFV--LNCMVPFCNPNE--GSDFFCISSTTFDVFVWFG 312

Qy        284 YSKAVADPFTYSL---LRRPFRQVLAGMVHRL 312
              :: :  :|  |:     |: |  :| |  |||
```

```
Db        313 WANSSLNPIIYAFNAGFRKAF-SILLG-CHRL 342


RESULT 14
Q9PVG0
ID   Q9PVG0      PRELIMINARY;      PRT;   367 AA.
AC   Q9PVG0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Somatostatin receptor type 1 subtype A.
GN   SST1A.
OS   Carassius auratus (Goldfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Carassius.
OX   NCBI_TaxID=7957;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20005543; PubMed=10537151;
RA   Lin X., Janovick J.A., Brothers S., Conn P.M., Peter R.E.;
RT   "Molecular cloning and expression of two type one somatostatin
RT   receptors in goldfish brain.";
RL   Endocrinology 140:5211-5219(1999).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AF097726; AAF08613.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   367 AA;  41658 MW;  BE85E6C26C103614 CRC64;

  Query Match             12.2%;  Score 227.5;  DB 13;  Length 367;
  Best Local Similarity   23.0%;  Pred. No. 5e-11;
  Matches   86;  Conservative   75;  Mismatches  160;  Indels   53;  Gaps   14;

Qy          2 GPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMP 61
              |      :: :  :|  | |  |::|:       |:::| |: : ::||::   ||  | :|
Db         31 GSSAIFISFIYSVVCLVGLCGNSMVIYVIFRYAKMKT-ATNIYILNLAIADELL-MLSVP 88

Qy         62 FTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPR 121
              | :   :    |      |:::  :|       :   :  || |::::|  |:: |   ||
Db         89 FLVTSSLLHHWPFGSLLCRLVLSVDAINMFTSIYCLTVLSIDRYISVVHPIKAARYRRPT 148

Qy        122 YAGLLLGCAWGQS-------LAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTA 174
              | ::    |  |       : ||  |      |   :|:::: |||||   | |
Db        149 IAKMVNLGVWMFSILVILPIIIFSTTAPN------SDGSVACNMQM-PEPERQWMAVFVI 201

Qy        175 TLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKR 234
                :||: |:  :|:              |  :  | |: :|| |             |:|:
Db        202 YAFLMGFLFPVIAICM-----------CYILIIVKMRVVALKAGWQ---------QRKK 240

Qy        235 RRHRATRKIGIAIATFLICFAPYVMTRLAELVPFVTV-NAQWGILSKCLTYSKAVADPFT 293
                : |  : : :  |:||: |: : :|  :  ||   ||    |:  | |: : |:|
Db        241 SERKITLMVMMVVTVFVICWMPFHIMQLVSV--FVQQHNATLSQLAVILGYANSCANPIL 298

Qy        294 YSLL----RRPFRQVLAGMVHRLLKRTPRPASTHDSSLDVAG-MVHQLLKRTPRPASTHN 348
              |  |    || |:::|   :      |  |   : ::|   |  |       :| :  :
Db        299 YGFLSDNFRRSFQRILC--LRWWDNATEEPIDYYATALKSRGYSVDDF-----QPENLES 351

Qy        349 GSVDTENDSCLQQT 362
              ||    | :|  :|
Db        352 GST-YRNGTCTSRT 364


RESULT 15
O96716
ID   O96716      PRELIMINARY;      PRT;   391 AA.
AC   O96716;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Dopamine D1/beta receptor.
OS   Branchiostoma lanceolatum (Common lancelet) (Amphioxus).
OC   Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC   Branchiostoma.
OX   NCBI_TaxID=7740;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Cardinaud B., Gibert JM., Sugamori K.S., Coudouel S., Guibert B.,
RA   Vincent J.D., Niznik H.B., Vernier P.;
RT   "The amphioxus D1/beta receptor and the emergence of the vertebrate
RT   adrenergic system.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR   EMBL; AJ005433; CAA06536.1; -.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 1.
DR   PRINTS; PR00237; GPCRRHODOPSN.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR   PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW   G-protein coupled receptor; Receptor; Transmembrane.
SQ   SEQUENCE   391 AA;  42622 MW;  67A5AD944AFA3FBE CRC64;

  Query Match             12.2%;  Score 227.5;  DB 5;  Length 391;
  Best Local Similarity   20.9%;  Pred. No. 5.3e-11;
  Matches   77;  Conservative   80;  Mismatches  159;  Indels   53;  Gaps   12;

Qy          5 EALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTL 64
              :|:|  : |::: : :: | ||:|      ::||  :   :|:|:   | :    :||
Db         39 QAVLGLITVILILLTVIGNVLVILAVTCHRKMRT-VTNFFIVSLACADLSVGITVLPFAA 97

Qy         65 LGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAG 124
                : |  |    | | |      |   :: : |::  :: |::||:  |  |  |:  | ||
Db         98 TNDILGYWPFG-GYCDVWVSFDVLNSTASILNLVVIAFDRFLAITAPFTYHTRMTERTAG 156

Qy        125 LLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRF---AAFTATLHAVGF 181
              :|:   || ||  |   :   | |     :  :| :  :|     |     |:|     : |
Db        157 ILIATVWGISLVVSFLPIQAGW--YRDNQSEEALAIYSDPCLCIFTASTAYTIVSSLISF 214

Qy        182 VLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRR----- 236
              :|| :: :     : : ||   :::: :  :                    ::|:  |
Db        215 YIPLLIMLVFYGIIFKAARDQARKINALEGR----------------LEQENNRGKKIS 257

Qy        237 ----HRATRKIGIAIATFLICFAPYVMTRLAE------LVPFVTVNAQW-GILSKCLTYS 285
                   :| : :|| :  |::|: |: : :          : | | :  | | :: |
Db        258 LAKEKKAAKTLGIIMGVFILCWLPFFVVNIVNPFCDRCVQPAVFIALTWLGWINSCF--- 314

Qy        286 KAVADPFTYSLLRRPFRQVLAGMV--HRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRP 343
                :|  |:   : ||:|   |:  |:    |  |   : : |:   :  :| ||
Db        315 ----NPIIYA-FNKEFRKVFVKMICCHKCRGVTVGP---NHADLNYDPVAMRLKKRGENA 366

Qy        344 ASTHNGSVD 352
              | ||  :
Db        367 NGTVNGDAN 375


Search completed: May 18, 2004, 12:08:08
Job time : 48 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 18, 2004, 12:06:02 ; Search time 22 Seconds
(without alignments)
851.829 Million cell updates/sec

Title: US-10-049-569-2
Perfect score: 1858
Sequence: 1 MGPGEALLAGLLVMVLAVAL.....ASTHNGSVDTENDSCLQQTH 363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	848	45.6	337	2	US-09-013-634-2	Sequence 7, Appli
2	236	12.7	348	3	US-08-875-540-13	Sequence 13, Appl
3	236	12.7	348	4	US-09-171-456-17	Sequence 17, Appl
4	236	12.7	348	4	US-09-473-634-13	Sequence 13, Appl
5	232.5	12.5	359	2	US-08-748-485-5	Sequence 5, Appli
6	232.5	12.5	391	3	US-08-120-601B-8	Sequence 8, Appli
7	230.5	12.4	402	1	US-08-444-734A-6	Sequence 6, Appli
8	230.5	12.4	402	1	US-08-087-772A-15	Sequence 15, Appl
9	230.5	12.4	402	3	US-08-846-704-4	Sequence 4, Appli
10	230.5	12.4	408	1	US-07-916-901-2	Sequence 2, Appli
11	230.5	12.4	408	3	US-08-450-962-2	Sequence 2, Appli
12	230.5	12.4	408	3	US-08-450-962-5	Sequence 5, Appli
13	230.5	12.4	408	4	US-08-848-631-2	Sequence 2, Appli
14	230.5	12.4	408	4	US-08-848-631-5	Sequence 5, Appli
15	230.5	12.4	425	4	US-09-479-128-2	Sequence 2, Appli
16	229.5	12.4	402	4	US-08-462-509B-2	Sequence 2, Appli
17	229.5	12.4	402	5	PCT-US95-05616-2	Sequence 2, Appli
18	229.5	12.4	425	3	US-08-846-704-2	Sequence 2, Appli
19	229.5	12.4	425	4	US-09-211-823C-22	Sequence 22, Appl
20	228.5	12.3	391	1	US-07-816-283-2	Sequence 2, Appli
21	228.5	12.3	391	1	US-07-816-283-4	Sequence 4, Appli
22	228.5	12.3	391	1	US-08-417-103-2	Sequence 2, Appli
23	228.5	12.3	391	1	US-08-417-103-4	Sequence 4, Appli
24	228.5	12.3	391	1	US-08-417-103-14	Sequence 14, Appl
25	228.5	12.3	477	1	US-07-791-936A-2	Sequence 2, Appli
26	228.5	12.3	477	1	US-08-383-781B-2	Sequence 2, Appli
27	228.5	12.3	477	2	US-07-969-267B-2	Sequence 2, Appli
28	228.5	12.3	477	4	US-09-168-510-2	Sequence 2, Appli
29	228	12.3	385	4	US-08-981-700A-4	Sequence 4, Appli
30	228	12.3	387	3	US-08-993-088A-7	Sequence 7, Appli
31	228	12.3	387	4	US-08-993-424B-7	Sequence 7, Appli
32	228	12.3	387	4	US-09-595-549-10	Sequence 10, Appl
33	228	12.3	387	4	US-09-603-680-7	Sequence 7, Appli
34	228	12.3	387	4	US-08-899-112B-28	Sequence 28, Appl
35	228	12.3	475	1	US-07-686-591-4	Sequence 4, Appli
36	228	12.3	475	1	US-07-970-715-4	Sequence 4, Appli
37	227.5	12.2	408	1	US-08-351-473B-3	Sequence 3, Appli
38	226.5	12.2	359	2	US-08-103-170-4	Sequence 4, Appli
39	226.5	12.2	359	2	US-08-103-170-6	Sequence 6, Appli
40	226.5	12.2	400	2	US-08-103-170-9	Sequence 9, Appli
41	226	12.2	353	2	US-08-467-559B-10	Sequence 10, Appl
42	225.5	12.1	365	2	US-08-428-243-9	Sequence 9, Appli
43	225.5	12.1	365	5	PCT-US93-10301-9	Sequence 9, Appli
44	224	12.1	369	3	US-08-120-601B-9	Sequence 9, Appli
45	224	12.1	372	2	US-08-626-685A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-013-634-2
; Sequence 7, Application US/08991946A
; Patent No. 5945306
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Tang, Tom Y.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: RAS PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,946A
; FILING DATE: December 16, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0445 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; Sequence 2, Application US/09013634
; Patent No. 5945307
; GENERAL INFORMATION:
; APPLICANT: M. Alexandra Glucksmann and Keith Robison

```
;    TITLE OF INVENTION:  LIGAND RECEPTORS AND USES THEREFOR
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/013,634
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Attorney, Jean M. Silveri
;      REGISTRATION NUMBER:  39,030
;      REFERENCE/DOCKET NUMBER:  MNI-036
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  337 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-013-634-2

  Query Match             45.6%;  Score 848;  DB 2;  Length 337;
  Best Local Similarity   52.0%;  Pred. No. 9.6e-66;
  Matches  170;  Conservative   52;  Mismatches  101;  Indels    4;  Gaps    1;

Qy          1 MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              |   :| ||||||  : |:|||||||||||  :||::| :|  :  :||: |:||   ::|
Db          1 MNSWDAGLAGLLVGTMGVSLLSNALVLLCLLHSADIRRQAPALFTLNLTCGNLLCTVVNM 60

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              | || ||:  | |:    |::  |||||||:|: ||:|||| |:|:|| ||| |  ::|
Db         61 PLTLAGVVARRQPAGDRLCRLAAFLDTFLAANSMLSMAALSIDRWVAVVFPLSYRAKMRL 120

Qy        121 RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG 180
              | | |::   |  :| |  |||  ||||:   :|||:|       || ||| ||   ||:
Db        121 RDAALMVAYTWLHALTFPAAALALSWLGFHQLYASCTLCSRRPDERLRFAVFTGAFHALS 180

Qy        181 FVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRAT 240
              |:|   ||| | :| :||| ||:|:| :||: | || |||||||:||| :|||| |||
Db        181 FLLSFVVLCCTYLKVLKVARFHCKRIDVITMQTLVLLVDLHPSVRERCLEEQKRRRQRAT 240

Qy        241 RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP 300
              :||   | |||:||||||:||| ||   | : : ||:|||| |||| :||| |||||
Db        241 KKISTFIGTFLVCFAPYVITRLVELFSTVPIGSHWGVLSKCLAYSKAASDPFVYSLLRHQ 300

Qy        301 FRQVLAGMVHRLLKRTPRPASTHDSSL 327
              :|:    :::||| |      | | | |
Db        301 YRKSCKEILNRLLHR----RSIHSSGL 323


RESULT 2
US-08-875-540-13
; Sequence 13, Application US/08875540A
; Patent No. 6015888
; GENERAL INFORMATION:
;  APPLICANT: Heath, Paul Roy
;  APPLICANT: Orange, Paul Richard
;  APPLICANT: Pearson, Ronald Carl Alan
;  APPLICANT: Wright, Simon Ralph
;  TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO THE DETECTION OF VARIATIONS IN
;  TITLE OF INVENTION: HUMAN HISTAMINE H2 RECEPTORS
;  FILE REFERENCE: 09347/002001
;  CURRENT APPLICATION NUMBER: US/08/875,540A
;  CURRENT FILING DATE: 1998-01-05
;  EARLIER APPLICATION NUMBER: PCT/EP96/00397
;  EARLIER FILING DATE: 1997-01-30
;  EARLIER APPLICATION NUMBER: GB9503866.7
;  EARLIER FILING DATE: 1995-01-30
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-875-540-13

  Query Match             12.7%;  Score 236;  DB 3;  Length 348;
  Best Local Similarity   24.1%;  Pred. No. 9.8e-13;
  Matches   85;  Conservative   60;  Mismatches  146;  Indels   62;  Gaps   11;

Qy         12 LVMVLAVALL----SNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGV 67
              : :|||| :|      | :| |    :  ||   :  :|:|::  |||  | :||: :
Db         18 ITVVLAVLILITVAGNVVVCLAVGLNRRLR-NLTNCFIVSLAITDLLLGLLVLPFSAIYQ 76

Qy         68 MRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLLL 127
              :  :       | :   ||  | : : |::  :| |:: ||  ||||   : |    : |
Db         77 LSCKWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMDPLRYPVLVTPARVAISL 136

Qy        128 GCAWGQSLAFSGAALGCSWLGYSS------AFASCSLRLPPEPERPRFAAFTATLHAVGF 181
              |  |: |  ::   |  :          : |::::          :        | |
Db        137 VLIWVISITLSFLSIHLGWNSRNETSKGNHTTSKCNVQVN--------EVYGLVDGLVTF 188

Qy        182 VLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRATR 241
              ||| ::|:|  :: ||||   :|:| ::    | :                | ||||
Db        189 YLPLLIMCITYYRIFRVARDQAKRIDHISSWKAATI----------------REHRATV 231

Qy        242 KIGIAIATFLICFAPYVMTRLAELVPFV--------TVNAQWGILSKCLTYSKAVADPFT 293
              :   :  |:||: ||          ||          :|     :   | |: :  :|
Db        232 TLAAVMGAFIICWFPY-------FTAFVYRGLRGDDAINEMLEAIVLWLGYANSALNPIL 284

Qy        294 YSLLRRPFR---QVLAGMVHRLLKRTPRPASTHDSSL-DVAGMVHQLLKRTPR 342
              |: | | ||   | |      || |      ::| :||   |  : :   | ||
Db        285 YAALNRDFRTGYQQL--FCCRLANR-----NSHKTSLRSNASQLSRTQSREPR 330


RESULT 3
US-09-171-456-17
; Sequence 17, Application US/09171456A
; Patent No. 6346380
; GENERAL INFORMATION:
;  APPLICANT: Wright, Simon Ralph
;  APPLICANT: Heath, Paul Roy
;  APPLICANT: Orange, Paul Richard
;  APPLICANT: Pearson, Ronald Carl Alan
;  TITLE OF INVENTION: DETECTION OF VARIATIONS IN HUMAN H2 RECEPTORS
;  FILE REFERENCE: 09347/004001
;  CURRENT APPLICATION NUMBER: US/09/171,456A
;  CURRENT FILING DATE: 1999-08-03
;  EARLIER APPLICATION NUMBER: PCT/GB97/01075
;  EARLIER FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-171-456-17

  Query Match             12.7%;  Score 236;  DB 4;  Length 348;
  Best Local Similarity   24.1%;  Pred. No. 9.8e-13;
  Matches   85;  Conservative   60;  Mismatches  146;  Indels   62;  Gaps   11;
```

```
Qy          12 LVMVLAVALL----SNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGV 67
               : :|||| :|     | :| |    :  ||   :   :|:|::  |||  | :||: :
Db          18 ITVVLAVLILITVAGNVVVCLAVGLNRRLR-NLTNCFIVSLAITDLLLGLLVLPFSAIYQ 76

Qy          68 MRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLLL 127
               :  :       | :   ||  | : : |::  :| |:: ||  ||||   : |    : |
Db          77 LSCKWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMDPLRYPVLVTPARVAISL 136

Qy         128 GCAWGQSLAFSGAALGCSWLGYSS------AFASCSLRLPPEPERPRFAAFTATLHAVGF 181
                  |  |:  |  ::   |  :         : |::::          :      | |
Db         137 VLIWVISITLSFLSIHLGWNSRNETSKGNHTTSKCNVQVN--------EVYGLVDGLVTF 188

Qy         182 VLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRATR 241
                |||  ::|:|  :: ||||   :|:| ::    | :                | ||||
Db         189 YLPLLIMCITYYRIFRVARDQAKRIDHISSWKAATI----------------REHRATV 231

Qy         242 KIGIAIATFLICFAPYVMTRLAELVPFV--------TVNAQWGILSKCLTYSKAVADPFT 293
               :   :  |:||: ||           ||         :|    :   | |: :  :|
Db         232 TLAAVMGAFIICWFPY-------FTAFVYRGLRGDDAINEMLEAIVLWLGYANSALNPIL 284

Qy         294 YSLLRRPFR---QVLAGMVHRLLKRTPRPASTHDSSL-DVAGMVHQLLKRTPR 342
               |: | | ||   | |      ||  |      ::| :||   |  : :   | ||
Db         285 YAALNRDFRTGYQQL--FCCRLANR-----NSHKTSLRSNASQLSRTQSREPR 330


RESULT 4
US-09-473-634-13
; Sequence 13, Application US/09473634
; Patent No. 6440670
; GENERAL INFORMATION:
;  APPLICANT: Heath, Paul Roy
;  APPLICANT:  Orange, Paul Richard
;  APPLICANT:  Pearson, Ronald Carl Alan
;  APPLICANT:  Wright, Simon Ralph
;  TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO THE DETECTION OF VARIATIONS IN
;  TITLE OF INVENTION:  HUMAN HISTAMINE H2 RECEPTORS
;  FILE REFERENCE: 09347/002001
;  CURRENT APPLICATION NUMBER: US/09/473,634
;  CURRENT FILING DATE:  1999-12-28
;  PRIOR APPLICATION NUMBER: US 08/875,540
;  PRIOR FILING DATE: 1998-01-05
;  PRIOR APPLICATION NUMBER: PCT/EP96/00397
;  PRIOR FILING DATE: 1997-01-30
;  PRIOR APPLICATION NUMBER: GB9503866.7
;  PRIOR FILING DATE: 1995-01-30
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13
;   LENGTH: 348
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-473-634-13

  Query Match             12.7%;  Score 236;  DB 4;  Length 348;
  Best Local Similarity   24.1%;  Pred. No. 9.8e-13;
  Matches   85;  Conservative   60;  Mismatches  146;  Indels   62;  Gaps   11;

Qy          12 LVMVLAVALL----SNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGV 67
               : :|||| :|     | :| |    :  ||   :   :|:|::  |||  | :||: :
Db          18 ITVVLAVLILITVAGNVVVCLAVGLNRRLR-NLTNCFIVSLAITDLLLGLLVLPFSAIYQ 76

Qy          68 MRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLLL 127
               :  :       | :   ||  | : : |::  :| |:: ||  ||||   : |    : |
Db          77 LSCKWSFGKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMDPLRYPVLVTPARVAISL 136

Qy         128 GCAWGQSLAFSGAALGCSWLGYSS------AFASCSLRLPPEPERPRFAAFTATLHAVGF 181
                  |  |:  |  ::   |  :         : |::::          :      | |
Db         137 VLIWVISITLSFLSIHLGWNSRNETSKGNHTTSKCNVQVN--------EVYGLVDGLVTF 188

Qy         182 VLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQKRRRHRATR 241
                |||  ::|:|  :: ||||   :|:| ::    | :                | ||||
Db         189 YLPLLIMCITYYRIFRVARDQAKRIDHISSWKAATI----------------REHRATV 231

Qy         242 KIGIAIATFLICFAPYVMTRLAELVPFV--------TVNAQWGILSKCLTYSKAVADPFT 293
               :   :  |:||: ||           ||         :|    :   | |: :  :|
Db         232 TLAAVMGAFIICWFPY-------FTAFVYRGLRGDDAINEMLEAIVLWLGYANSALNPIL 284

Qy         294 YSLLRRPFR---QVLAGMVHRLLKRTPRPASTHDSSL-DVAGMVHQLLKRTPR 342
               |: | | ||   | |      ||  |      ::| :||   |  : :   | ||
Db         285 YAALNRDFRTGYQQL--FCCRLANR-----NSHKTSLRSNASQLSRTQSREPR 330


RESULT 5
US-08-748-485-5
; Sequence 5, Application US/08748485
; Patent No. 5817480
;  GENERAL INFORMATION:
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Murry, Lynn E.
;    TITLE OF INVENTION:  NOVEL HISTAMINE H2 RECEPTOR
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/748,485
;      FILING DATE:  Herewith
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0159 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  791239
US-08-748-485-5

  Query Match             12.5%;  Score 232.5;  DB 2;  Length 359;
  Best Local Similarity   22.0%;  Pred. No. 2e-12;
  Matches   73;  Conservative   67;  Mismatches  139;  Indels   53;  Gaps    8;

Qy          11 LLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGVMRG 70
               :|:::: | :  | :| |     :  ||:  :   :|:|::  |||  | :||: :  :
Db          24 ILIILILVTVAGNVVVCLAVGLNRRLRS-LTNCFIVSLAVTDLLLGLLVLPFSAIYQLSC 82

Qy          71 RTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLLLGCA 130
               :   :   | :   ||  | : : |::  :| |:: ||  ||||   : |    : |
Db          83 KWSFSKVFCNIYTSLDVMLCTASILNLFMISLDRYCAVTDPLRYPVLITPARVAISLVFI 142

Qy         131 WGQSLAFSGAALGCSWLGY------SSAFASCSLRLPPEPERPRFAAFTATLHAVGFVLP 184
               |  |:  |  ::   |          :     | :::           :       | | ||
```

```
Db        143 WVISITLSFLSIHLGWNSRNETSKDNDTIVKCKVQVN--------EVYGLVDGLVTFYLP 194

Qy        185 LAVLCLTSLQVHRVARRHCQRMDTV-TMKALALLADLHPSVRQRCLIQQKRRRHRATRKI 243
              | ::|:|  :: ::||   :|:: : : ||  :                  | |:|| :
Db        195 LLIMCITYFRIFKIAREQARRINHIGSWKAATI----------------REHKATVTL 236

Qy        244 GIAIATFLICFAPYVMTRLAELVPPFV--------TVNAQWGILSKCLTYSKAVADPFTYS 295
                :  |:||: ||          ||          ||  :  :   | |: :  :|  |:
Db        237 AAVMGAFIICWFPY-------FTVFVYRGLKGDDAVNEVFEDVVLWLGYANSALNPILYA 289

Qy        296 LLRRPFRQVLAGMVHRLLKRTPRPASTHDSSL 327
                | | ||       |:|       ::|::||
Db        290 ALNRDFRTA----YHQLFCCRLASHNSHETSL 317


RESULT 6
US-08-120-601B-8
; Sequence 8, Application US/08120601B
; Patent No. 6235496
;  GENERAL INFORMATION:
;    APPLICANT:  Yu, Lei
;    TITLE OF INVENTION:  MU OPIOID RECEPTORS:  COMPOSOTIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/120,601B
;      FILING DATE:  13-SEP-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wilson, Mark B.
;      REGISTRATION NUMBER:  37,259
;      REFERENCE/DOCKET NUMBER:  INDA:002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  391 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-120-601B-8

  Query Match            12.5%;  Score 232.5;  DB 3;  Length 391;
  Best Local Similarity   24.5%;  Pred. No. 2.3e-12;
  Matches   78;  Conservative   62;  Mismatches  130;  Indels   49;  Gaps   11;

Qy          2 GPGEALLAGLLVMVLA-VALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDM 60
              | | |:|   :  |:  | |  |::|:      |:::| |: : ::||::   ||  | :
Db         54 GQGSAILISFIYSVVCLVGLCGNSMVIYVILRYAKMKT-ATNIYILNLAIADELL-MLSV 111

Qy         61 PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRP 120
              || :   :   |    |:::  :| :   :   :   || |:::||  |:: |   ||
Db        112 PFLVTSTLLRHWPFGALLCRLVLSVDAYNMFTSIYCLTVLSVDRYVAVVHPIKAARYRRP 171

Qy        121 RYAGLLLGCAWGQSL-------AFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFT 173
                | ::    |  ||         ||  |         |    :|:: | ||| :      |
Db        172 TVAKVVNLGVWVLSLLVILPIVVFSRTAAN------SDGTVACNM-LMPEPAQRWLVGFV 224

Qy        174 ATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALALLADLHPSVRQRCLIQQK 233
                  :||:||:  :||           |  :    |: :|| |           |:|
```

```
Db        225 LYTFLMGFLLPVGAICL----------CYVLIIAKMRMVALKAGWQ---------QRK 263

Qy        234 RRRHRATRKIGIAIATFLICFAPYVMTRLAELV---PFVTVNAQWGILSKCLTYSKAVAD 290
              |   : |  : :  :|:|: |: : :|  :       ||:     ||  | |: : |:
Db        264 RSERKITLMVMMVVMVFVICWMPFYVVQLVNVFAEQDDATVSQ----LSVILGYANSCAN 319

Qy        291 PFTYSLL----RRPFRQVL 305
              |  |  |    :| |:::|
Db        320 PILYGFLSDNFKRSFQRIL 338


RESULT 7
US-08-444-734A-6
; Sequence 6, Application US/08444734A
; Patent No. 5610282
;  GENERAL INFORMATION:
;    APPLICANT:  Sibley, David R.
;    APPLICANT:  Monsma, Frederick J.
;    APPLICANT:  Mahan, Lawrence C.
;    APPLICANT:  McVittie, Loris D.
;    TITLE OF INVENTION:  cDNA encoding the rat D1 dopamine
;    TITLE OF INVENTION:  receptor linked to adnylyl cyclase activation and
;    TITLE OF INVENTION:  expression of the receptor protein in plasmid-transfected
;    TITLE OF INVENTION:  cell lines
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe, Martens, Olson and Bear
;      STREET:  620 Newport Center Drive, Sixteenth Floor
;      CITY:  Newport Beach
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/444,734A
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/029,917
;      FILING DATE:  03-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/548,714
;      FILING DATE:  06-JUL-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Altman, Daniel E.
;      REGISTRATION NUMBER:  34,115
;      REFERENCE/DOCKET NUMBER:  NIH065.001FW1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (714) 760-0404
;      TELEFAX:  (714) 760-9502
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  402 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  internal
US-08-444-734A-6

  Query Match            12.4%;  Score 230.5;  DB 1;  Length 402;
  Best Local Similarity   24.9%;  Pred. No. 3.5e-12;
  Matches   93;  Conservative   61;  Mismatches  154;  Indels   65;  Gaps   14;

Qy          3 PGEALLAGLLVMVLAVALLS----NALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAAL 58
              | || ||| |   ||:|:|:    | ||::  |::  |:|  : | : :|:   |::  |
Db         34 PWEAALAGAL---LALAVLATVGGNLLVIVAIAWTPRLQT-MTNVFVTSLAAADLVMGLL 89
```

```
Qy          59 DMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRL 118
Db          90 VVPPAATLALTGHWPLGATGCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALV 149

Qy         119 RPRYAGLLLGCAWGQSLAFSGAALGCSW--LGYSSAFASCSLRLPPEPERPRFAAFTATL 176
Db         150 TKRCARTAVVLVWVVSAAVSFAPIMSQWWRVGADAEAQRCH-------SNPRCCAFASNM 202

Qy         177 ------HAVGFVLPLAVLCLTSLQVHRVARRHC-----------QRMDTVTMKALA--- 215
Db         203 PYVLLSSSVSFYLPLLVMLFVYARVFVVATRQLRLLRGELGRFPPEESPPAPSRSLAPAP 262

Qy         216 ----LLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRL------AEL 265
Db         263 VGTCAPPEGVPACGRRPARLLPLREHRALCTLGLIMGTFTLCWLPFFLANVLRALGGPSL 322

Qy         266 VP---FVTVNAQWGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTPRP 319
Db         323 VPGPAFLALN--W------LGYANSAFNPLIYCRSPDFRSAFRRLLCRCGRRL---PPEP 371

Qy         320 ASTHDSSLDVAGM 332
Db         372 CAAARPALFPSGV 384
```

RESULT 8
US-08-087-772A-15
; Sequence 15, Application US/08087772A
; Patent No. 5691155
; GENERAL INFORMATION:
; APPLICANT: Nahmias, Clara
; APPLICANT: Emorine, Jean L.
; APPLICANT: Strosberg, Donny A.
; TITLE OF INVENTION: Nucleotide Sequences Encoding the Murine
; TITLE OF INVENTION: Beta3-Adrenergic Receptor and Their Applications
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5691155th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,772A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Linker, Raymond O.
; REGISTRATION NUMBER: 26,419
; REFERENCE/DOCKET NUMBER: 3339-195
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-087-772A-15

Query Match 12.4%; Score 230.5; DB 1; Length 402;
Best Local Similarity 24.9%; Pred. No. 3.5e-12;
Matches 93; Conservative 61; Mismatches 154; Indels 65; Gaps 14;

```
Qy           3 PGEALLAGLLVMVLAVALLS----NALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAAL 58
Db          34 PWEAALAGAL---LALAVLATVGGNLLVIVAIAWTPRLQT-MTNVFVTSLAAADLVMGLL 89

Qy          59 DMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRL 118
Db          90 VVPPAATLALTGHWPLGATGCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALV 149

Qy         119 RPRYAGLLLGCAWGQSLAFSGAALGCSW--LGYSSAFASCSLRLPPEPERPRFAAFTATL 176
Db         150 TKRCARTAVVLVWVVSAAVSFAPIMSQWWRVGADAEAQRCH-------SNPRCCAFASNM 202

Qy         177 ------HAVGFVLPLAVLCLTSLQVHRVARRHC-----------QRMDTVTMKALA--- 215
Db         203 PYVLLSSSVSFYLPLLVMLFVYARVFVVATRQLRLLRGELGRFPPEESPPAPSRSLAPAP 262

Qy         216 ----LLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRL------AEL 265
Db         263 VGTCAPPEGVPACGRRPARLLPLREHRALCTLGLIMGTFTLCWLPFFLANVLRALGGPSL 322

Qy         266 VP---FVTVNAQWGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTPRP 319
Db         323 VPGPAFLALN--W------LGYANSAFNPLIYCRSPDFRSAFRRLLCRCGRRL---PPEP 371

Qy         320 ASTHDSSLDVAGM 332
Db         372 CAAARPALFPSGV 384
```

RESULT 9
US-08-846-704-4
; Sequence 4, Application US/08846704
; Patent No. 6020157
; GENERAL INFORMATION:
; APPLICANT: BERGSMA, DERK J.
; APPLICANT: ELLIS, CATHERINE E.
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/846,704
; FILING DATE: 30-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-846-704-4

```
  Query Match             12.4%;  Score 230.5;  DB 3;  Length 402;
  Best Local Similarity   25.3%;  Pred. No. 3.5e-12;
  Matches   94;  Conservative   55;  Mismatches  154;  Indels   69;  Gaps   13;

Qy          5 EALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTL 64
              | :|     | |  |||: | || |     :  :||  :   :|||||  :|: |: :| :|
Db         46 EWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRT-VTNYFIVNLSLADVLVTAICLPASL 104

Qy         65 LGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAG 124
              |  :           |:|| :|     | | |::: :: |:| |:  || :    | |  |
Db        105 LVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTAR-RARG 163

Qy        125 LLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAA----------FTA 174
               :||  |  |||           :  :|   ||  ||    | |  :          :
Db        164 SILG-IWAVSLAI---------MVPQAAVMECSSVLPELANRTRLFSVCDERWADDLYPK 213

Qy        175 TLHAVGFVLP-LAVLCLTSLQVHRVARRHCQRMDTVTMKALA--------LLADL----- 220
               |:  |::  || | | ::   :: |:   |     |  ||             | ||
Db        214 IYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLS 273

Qy        221 -HPSVRQRCL---IQQKRRRHRATRKIGIAIATFLICFAP----------YVMTRLA--- 263
                |  | |      ::| | | :  : : : :  | :|: |           : | | |
Db        274 GEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDR 333

Qy        264 ELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRPFRQVL----------AGMVHRLL 313
               | |      : |       | |: : |:|  |: |   ||:            |   |
Db        334 EAVYACFTFSHW------LVYANSAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLK 387

Qy        314 KRTPRPASTHDS 325
               :|| :::| |
Db        388 APSPRSSASHKS 399


RESULT 10
US-07-916-901-2
; Sequence 2, Application US/07916901
; Patent No. 5364772
;  GENERAL INFORMATION:
;    APPLICANT:  Granneman, James G.
;    APPLICANT:  Lahners, Kristine N.
;    APPLICANT:  Rao, Donald D.
;    TITLE OF INVENTION:  @  @3-ADRENERGIC RECEPTOR PROTEIN AND DNA
;    TITLE OF INVENTION:  ENCODING SAME
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REISING, ETHINGTON, BARNARD, PERRY &
;      ADDRESSEE:  MILTON
;      STREET:  201 W. Big Beaver - Ste. 400; P.O. Box 4390
;      CITY:  Troy
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  48099
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/916,901
;      FILING DATE:  19920720
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kohn, Kenneth I.
;      REGISTRATION NUMBER:  30,955
;      REFERENCE/DOCKET NUMBER:  P-324(WSU)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (313) 689-3554
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  408 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-916-901-2

  Query Match             12.4%;  Score 230.5;  DB 1;  Length 408;
  Best Local Similarity   24.9%;  Pred. No. 3.5e-12;
  Matches   93;  Conservative   61;  Mismatches  154;  Indels   65;  Gaps   14;

Qy          3 PGEALLAGLLVMVLAVALLS----NALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAAL 58
              | ||  ||| |   ||:|:|:    | ||::  |::  :|  : | : :|:   |::  |
Db         34 PWEAALAGAL---LALAVLATVGGNLLVIVAIAWTPRLQT-MTNVFVTSLAAADLVMGLL 89

Qy         59 DMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRL 118
               :|      : |  |     |::   :|    : :  :: ||: |::|||  ||||   :
Db         90 VVPPAATLALTGHWPLGATGCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALV 149

Qy        119 RPRYAGLLLGCAWGQSLAFSGAALGCSW--LGYSSAFASCSLRLPPEPERPRFAAFTATL 176
               | |   :  |  | | | | :   |  :|  :    |          ||  || : :
Db        150 TKRCARTAVVLVWVVSAAVSFAPIMSQWWRVGADAEAQRCH-------SNPRCCAFASNM 202

Qy        177 ------HAVGFVLPLAVLCLTSLQVHRVARRHC------------QRMDTVTMKALA--- 215
                   :| | |||  |:      :|  || |                :       ::||
Db        203 PYVLLSSSVSFYLPLLVMLFVYARVFVVATRQLRLLRGELGRPPPEESPPAPSRSLAPAP 262

Qy        216 ----LLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRL------AEL 265
                   :  |:  :|       | |||   :|: :  || :|: |: :  :         |
Db        263 VGTCAPPEGVPACGRRPARLLPLREHRALCTLGLIMGTFTLCWLPFFLANVLRALGGPSL 322

Qy        266 VP---FVTVNAQWGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTPRP 319
               ||   |: :|  |       | |: :  :|  |       |  ||::|    ||    | |
Db        323 VPGPAFLALN--W------LGYANSAFNPLIYCRSPDFRSAFRRLLCRCGRRL---PPEP 371

Qy        320 ASTHDSSLDVAGM 332
               :    :|  :|:
Db        372 CAAARPALFPSGV 384


RESULT 11
US-08-450-962-2
; Sequence 2, Application US/08450962
; Patent No. 6274706
;  GENERAL INFORMATION:
;    APPLICANT:  EMORINE, Laurent; MARULLO, Stefano;
;    APPLICANT:  STROSBERG, Donny
;    TITLE OF INVENTION:  INTRON/EXON OF THE HUMAN AND
;    TITLE OF INVENTION:  GENES
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  KECK, MAHIN & CATE
;      STREET:  P.O. BOX 06110
;      CITY:  CHICAGO
;      STATE:  ILLINOIS
;      COUNTRY:  U.S.A.
;      ZIP:  60606-0110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3-1/2" diskette
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/450,962
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/117,829
;      FILING DATE:  08-SEPT-1993
;      APPLICATION NUMBER:  07/721,571
;      FILING DATE:  25-MAY-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR89/00918
;      FILING DATE:  25-JAN-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fleit, Martin; Gollin, Michael A.
;      REGISTRATION NUMBER:  16,900; 31,957
```

; REFERENCE/DOCKET NUMBER: 47078-042
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 789-3400
; TELEFAX: (202) 789-1158
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 residues
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: polypeptide
US-08-450-962-2

Query Match 12.4%; Score 230.5; DB 3; Length 408;
Best Local Similarity 24.9%; Pred. No. 3.5e-12;
Matches 93; Conservative 61; Mismatches 154; Indels 65; Gaps 14;

```
Qy          3 PGEALLAGLLVMVLAVALLS----NALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAAL 58
              | || ||| |   ||:|:|:    | ||::  |::  |:|  : | : :|:   |::  |
Db         34 PWEAALAGAL---LALAVLATVGGNLLVIVAIAWTPRLQT-MTNVFVTSLAAADLVMGLL 89

Qy         59 DMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRL 118
              :|      : |   |     |::   :|    : :  :: ||: |::|||  ||||   :
Db         90 VVPPAATLALTGHWPLGATGCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALV 149

Qy        119 RPRYAGLLLGCAWGQSLAFSGAALGCSW--LGYSSAFASCSLRLPPEPERPRFAAFTATL 176
               | |   :   |  | | | | | :   |  :|  :    |           ||  || : :
Db        150 TKRCARTAVVLVWVVSAAVSFAPIMSQWWRVGADAEAQRCH-------SNPRCCAFASNM 202

Qy        177 ------HAVGFVLPLAVLCLTSLQVHRVARRHC------------QRMDTVTMKALA--- 215
                   :| | ||| |:     :|  || |                :         ::||
Db        203 PYVLLSSSVSFYLPLLVMLFVYARVFVVATRQLRLLRGELGRFPPEESPPAPSRSLAPAP 262

Qy        216 ----LLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRL------AEL 265
                  :  |:  :|       | |||   :|: : || :|: |: :  :         |
Db        263 VGTCAPPEGVPACGRRPARLLPLREHRALCTLGLIMGTFTLCWLPFFLANVLRALGGPSL 322

Qy        266 VP---FVTVNAQWGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTPRP 319
              ||   |: :|  |       | |: :  :|  |       |  ||::|     ||    | |
Db        323 VPGPAFLALN--W------LGYANSAFNPLIYCRSPDFRSAFRRLLCRCGRRL---PPEP 371

Qy        320 ASTHDSSLDVAGM 332
               :    :|  :|:
Db        372 CAAARPALFPSGV 384
```

RESULT 12
US-08-450-962-5
; Sequence 5, Application US/08450962
; Patent No. 6274706
; GENERAL INFORMATION:
; APPLICANT: EMORINE, Laurent; MARULLO, Stefano;
; APPLICANT: STROSBERG, Donny
; TITLE OF INVENTION: INTRON/EXON OF THE HUMAN AND
; TITLE OF INVENTION: GENES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KECK, MAHIN & CATE
; STREET: P.O. BOX 06110
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: U.S.A.
; ZIP: 60606-0110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3-1/2" diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,962
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/117,829
; FILING DATE: 08-SEPT-1993
; APPLICATION NUMBER: 07/721,571
; FILING DATE: 25-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR89/00918
; FILING DATE: 25-JAN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Fleit, Martin; Gollin, Michael A.
; REGISTRATION NUMBER: 16,900; 31,957
; REFERENCE/DOCKET NUMBER: 47078-042
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 789-3400
; TELEFAX: (202) 789-1158
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 residues
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: polypeptide
US-08-450-962-5

Query Match 12.4%; Score 230.5; DB 3; Length 408;
Best Local Similarity 24.9%; Pred. No. 3.5e-12;
Matches 93; Conservative 61; Mismatches 154; Indels 65; Gaps 14;

```
Qy          3 PGEALLAGLLVMVLAVALLS----NALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAAL 58
              | || ||| |   ||:|:|:    | ||::  |::  |:|  : | : :|:   |::  |
Db         34 PWEAALAGAL---LALAVLATVGGNLLVIVAIAWTPRLQT-MTNVFVTSLAAADLVMGLL 89

Qy         59 DMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRL 118
              :|      : |   |     |::   :|    : :  :: ||: |::|||  ||||   :
Db         90 VVPPAATLALTGHWPLGATGCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALV 149

Qy        119 RPRYAGLLLGCAWGQSLAFSGAALGCSW--LGYSSAFASCSLRLPPEPERPRFAAFTATL 176
               | |   :   |  | | | | | :   |  :|  :    |           ||  || : :
Db        150 TKRCARTAVVLVWVVSAAVSFAPIMSQWWRVGADAEAQRCH-------SNPRCCAFASNM 202

Qy        177 ------HAVGFVLPLAVLCLTSLQVHRVARRHC------------QRMDTVTMKALA--- 215
                   :| | ||| |:     :|  || |                :         ::||
Db        203 PYVLLSSSVSFYLPLLVMLFVYARVFVVATRQLRLLRGELGRFPPEESPPAPSRSLAPAP 262

Qy        216 ----LLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRL------AEL 265
                  :  |:  :|       | |||   :|: : || :|: |: :  :         |
Db        263 VGTCAPPEGVPACGRRPARLLPLREHRALCTLGLIMGTFTLCWLPFFLANVLRALGGPSL 322

Qy        266 VP---FVTVNAQWGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTPRP 319
              ||   |: :|  |       | |: :  :|  |       |  ||::|     ||    | |
Db        323 VPGPAFLALN--W------LGYANSAFNPLIYCRSPDFRSAFRRLLCRCGRRL---PPEP 371

Qy        320 ASTHDSSLDVAGM 332
               :    :|  :|:
Db        372 CAAARPALFPSGV 384
```

RESULT 13
US-08-848-631-2
; Sequence 2, Application US/08848631
; Patent No. 6635442
; GENERAL INFORMATION:
; APPLICANT: EMORINE, Laurent; MARULLO, Stefano;
; STROSBERG, Donny
; TITLE OF INVENTION: INTRON/EXON OF THE HUMAN AND
; MOUSE a3-ADRENERGIC RECEPTOR
; GENES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KECK, MAHIN & CATE
; STREET: P.O. BOX 06110
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: U.S.A.

```
;            ZIP: 60606-0110
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3-1/2" diskette
;            COMPUTER: IBM compatible
;            OPERATING SYSTEM: MS-DOS
;            SOFTWARE: ASCII
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/848,631
;            FILING DATE: 08-Jun-1999
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 07/721,571
;            FILING DATE: 25-MAY-1990
;            APPLICATION NUMBER: PCT/FR89/00918
;            FILING DATE: 25-JAN-1989
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Fleit, Martin; Gollin, Michael A.
;            REGISTRATION NUMBER: 16,900; 31,957
;            REFERENCE/DOCKET NUMBER: 47078-042
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (202) 789-3400
;            TELEFAX: (202) 789-1158
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 408 residues
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: <Unknown>
;            DESCRIPTION: polypeptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-848-631-2

  Query Match             12.4%;  Score 230.5;  DB 4;  Length 408;
  Best Local Similarity   24.9%;  Pred. No. 3.5e-12;
  Matches   93;  Conservative   61;  Mismatches  154;  Indels   65;  Gaps   14;

Qy          3 PGEALLAGLLVMVLAVALLS----NALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAAL 58
              | || ||| |    ||:|:|:     | ||::  |::  |:|  : | : :|:    |::  |
Db         34 PWEAALAGAL---LALAVLATVGGNLLVIVAIAWTPRLQT-MTNVFVTSLAAADLVMGLL 89

Qy         59 DMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRL 118
               :|      : |  |      |::   :|    : :  :: ||: |::|||  ||||   :
Db         90 VVPPAATLALTGHWPLGATGCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALV 149

Qy        119 RPRYAGLLLGCAWGQSLAFSGAALGCSW--LGYSSAFASCSLRLPPEPERPRFAAFTATL 176
              |  |   :   |  | | | | | :   |  :|  :    |              ||  || : :
Db        150 TKRCARTAVVLVWVVSAAVSFAPIMSQWWRVGADAEAQRCH-------SNPRCCAFASNM 202

Qy        177 ------HAVGFVLPLAVLCLTSLQVHRVARRHC-----------QRMDTVTMKALA--- 215
                   :| | ||| |:     :|  || |              :         ::||
Db        203 PYVLLSSSVSFYLPLLVMLFVYARVFVVATRQLRLLRGELGRFPPEESPPAPSRSLAPAP 262

Qy        216 ----LLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRL------AEL 265
                  :  |:  :|       | |||   :|: : || :|: |: :  :          |
Db        263 VGTCAPPEGVPACGRRPARLLPLREHRALCTLGLIMGTFTLCWLPFFLANVLRALGGPSL 322

Qy        266 VP---FVTVNAQWGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTPRP 319
              ||   |: :|  |       | |: :  :|  |       |  ||::|     ||    | |
Db        323 VPGPAFLALN--W------LGYANSAFNPLIYCRSPDFRSAFRRLLCRCGRRL---PPEP 371

Qy        320 ASTHDSSLDVAGM 332
              :    :|  :|:
Db        372 CAAARPALFPSGV 384


RESULT 14
US-08-848-631-5
; Sequence 5, Application US/08848631
; Patent No. 6635442
;   GENERAL INFORMATION:
;       APPLICANT: EMORINE, Laurent; MARULLO, Stefano;
;                  STROSBERG, Donny
;       TITLE OF INVENTION: INTRON/EXON OF THE HUMAN AND
;                           MOUSE a3-ADRENERGIC RECEPTOR
;                   GENES
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: KECK, MAHIN & CATE
;            STREET: P.O. BOX 06110
;            CITY: CHICAGO
;            STATE: ILLINOIS
;            COUNTRY: U.S.A.
;            ZIP: 60606-0110
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3-1/2" diskette
;            COMPUTER: IBM compatible
;            OPERATING SYSTEM: MS-DOS
;            SOFTWARE: ASCII
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/848,631
;            FILING DATE: 08-Jun-1999
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 07/721,571
;            FILING DATE: 25-MAY-1990
;            APPLICATION NUMBER: PCT/FR89/00918
;            FILING DATE: 25-JAN-1989
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Fleit, Martin; Gollin, Michael A.
;            REGISTRATION NUMBER: 16,900; 31,957
;            REFERENCE/DOCKET NUMBER: 47078-042
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (202) 789-3400
;            TELEFAX: (202) 789-1158
;   INFORMATION FOR SEQ ID NO: 5:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 408 residues
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: <Unknown>
;            DESCRIPTION: polypeptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-08-848-631-5

  Query Match             12.4%;  Score 230.5;  DB 4;  Length 408;
  Best Local Similarity   24.9%;  Pred. No. 3.5e-12;
  Matches   93;  Conservative   61;  Mismatches  154;  Indels   65;  Gaps   14;

Qy          3 PGEALLAGLLVMVLAVALLS----NALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAAL 58
              | || ||| |    ||:|:|:     | ||::  |::  |:|  : | : :|:    |::  |
Db         34 PWEAALAGAL---LALAVLATVGGNLLVIVAIAWTPRLQT-MTNVFVTSLAAADLVMGLL 89

Qy         59 DMPFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRL 118
               :|      : |  |      |::   :|    : :  :: ||: |::|||  ||||   :
Db         90 VVPPAATLALTGHWPLGATGCELWTSVDVLCVTASIETLCALAVDRYLAVTNPLRYGALV 149

Qy        119 RPRYAGLLLGCAWGQSLAFSGAALGCSW--LGYSSAFASCSLRLPPEPERPRFAAFTATL 176
              |  |   :   |  | | | | | :   |  :|  :    |              ||  || : :
Db        150 TKRCARTAVVLVWVVSAAVSFAPIMSQWWRVGADAEAQRCH-------SNPRCCAFASNM 202

Qy        177 ------HAVGFVLPLAVLCLTSLQVHRVARRHC-----------QRMDTVTMKALA--- 215
                   :| | ||| |:     :|  || |              :         ::||
Db        203 PYVLLSSSVSFYLPLLVMLFVYARVFVVATRQLRLLRGELGRFPPEESPPAPSRSLAPAP 262

Qy        216 ----LLADLHPSVRQRCLIQQKRRRHRATRKIGIAIATFLICFAPYVMTRL------AEL 265
                  :  |:  :|       | |||   :|: : || :|: |: :  :          |
Db        263 VGTCAPPEGVPACGRRPARLLPLREHRALCTLGLIMGTFTLCWLPFFLANVLRALGGPSL 322

Qy        266 VP---FVTVNAQWGILSKCLTYSKAVADPFTYSL---LRRPFRQVLAGMVHRLLKRTPRP 319
              ||   |: :|  |       | |: :  :|  |       |  ||::|     ||    | |
Db        323 VPGPAFLALN--W------LGYANSAFNPLIYCRSPDFRSAFRRLLCRCGRRL---PPEP 371

Qy        320 ASTHDSSLDVAGM 332
              :    :|  :|:
Db        372 CAAARPALFPSGV 384


RESULT 15
```

```
US-09-479-128-2
; Sequence 2, Application US/09479128
; Patent No. 6319710
; GENERAL INFORMATION:
;  APPLICANT:  Berglind Ran Olafsdottir
;  APPLICANT:  Jeffrey Gulcher
;  TITLE OF INVENTION:  HUMAN NARCOLEPSY GENE
;  FILE REFERENCE:  2345.1005-001
;  CURRENT APPLICATION NUMBER: US/09/479,128
;  CURRENT FILING DATE:  2000-01-07
;  PRIOR APPLICATION NUMBER:  US 09/379,083
;  PRIOR FILING DATE:  1999-08-23
;  NUMBER OF SEQ ID NOS:  22
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 425
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-09-479-128-2

  Query Match             12.4%;  Score 230.5;  DB 4;  Length 425;
  Best Local Similarity   25.3%;  Pred. No. 3.7e-12;
  Matches   94;  Conservative   55;  Mismatches  154;  Indels   69;  Gaps   13;

Qy          5 EALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTL 64
              | :|     | |  |||: | || |     :  :||  :   :|||||  :|: |: :| :|
Db         46 EWVLIAAYVAVFVVALVGNTLVCLAVWRNHHMRT-VTNYFIVNLSLADVLVTAICLPASL 104

Qy         65 LGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAG 124
              |  :          |:|| :|     | | |::: :: |:| |:  || :    | |  |
Db        105 LVDITESWLFGHALCKVIPYLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTAR-RARG 163

Qy        125 LLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAA----------FTA 174
              :||  |  |||          :   :|   ||  ||    | |  :          :
Db        164 SILG-IWAVSLAI---------MVPQAAVMECSSVLPELANRTRLFSVCDERWADDLYPK 213

Qy        175 TLHAVGFVLP-LAVLCLTSLQVHRVARRHCQRMDTVTMKALA--------LLADL----- 220
              |:  |::  || | |  ::   :: |:   |     |  ||           | ||
Db        214 IYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIPGTTSALVRNWKRPSDQLGDLEQGLS 273

Qy        221 -HPSVRQRCL---IQQKRRRHRATRKIGIAIATFLICFAP----------YVMTRLA--- 263
               |  | |     ::| | | :  : : : :  | :|: |          : | | |
Db        274 GEPQPRARAFLAEVKQMRARRKTAKMLMVVLLVFALCYLPISVLNVLKRVFGMFRQASDR 333

Qy        264 ELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRPFRQVL----------AGMVHRLL 313
              | |      : |       | |: : |:|  |: |   ||:             |    |
Db        334 EAVYACFTFSHW------LVYANSAANPIIYNFLSGKFREQFKAAFSCCLPGLGPCGSLK 387

Qy        314 KRTPRPASTHDS 325
              :|| :::| |
Db        388 APSPRSSASHKS 399


Search completed: May 18, 2004, 12:09:16
Job time : 24 secs
```